BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

REMARKS

Claims 1-12 and 43-45 are pending in the instant application. Claims 13-42, and 46-58 have been cancelled without prejudice or disclaimer. Claims 1-4 and have been amended without prejudice to their refiling as originally filed in future continuation or divisional applications. Support for the amendments may be found throughout the specification, for example on page 23, line 31 through page 24, line 1, on page 24, lines 20 and 24, and on page 17, line 8. No new matter has been added as a result of the above-described amendments. The rejections set forth in the Office Action have been overcome by amendment or are traversed by argument below.

1. Specification

The Office Action asserts that the title of the application is not descriptive and thus requires a new title. The Applicants have amended the title to recite "Nucleic Acid Molecules Encoding Leucine-Rich Repeat-Containing G-Protein Coupled Receptor-8 Proteins and Uses Thereof," a title Applicants respectfully submit is clearly indicative of the invention to which the claims are directed.

2. Claim Objections

The Office Action has objected to Claims 1-3 for reciting unelected inventions. In response Applicants have amended the claims to reflect Applicant's sequence election. Applicants therefore respectfully request the objection to Claims 1-3 be withdrawn.

3. Claim Rejections – 35 U.S.C. § 112, second paragraph

(A) The Office Action has rejected Claims 1-3, and dependent Claims 9-12 and 43-45 under 35 U.S.C. § 112, second paragraph for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. Specifically, the Office Action asserts that the recitation of moderately and stringent conditions renders the claim indefinite.

Applicants have amended Claims 1-3 to recite "a nucleotide sequence that hybridizes to the complement of the [previously recited] nucleotide sequence[s] [] at 50°C in a hybridization buffer comprising 0.015 M sodium chloride and 0.0015 M sodium citrate." Support for the amendments may be found in the specification on page 23, line 31 through page 24, line 1. Applicants contend

that as amended, the Claims are no longer indefinite and respectfully request that this rejection be withdrawn.

(B) The Office Action has rejected Claim 8, and dependent Claims 9-10, under 35 U.S.C. § 112, second paragraph for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. Specifically, the Office Action asserts that it is not clear how the polynucleotide complements of Claims 1(d), 2(f), and 3(h) produce the polypeptide disclosed in the instant application.

Applicants traverse this assertion, and note that any vector comprising, for example, SEQ ID NO: 1 would necessarily comprise the sequence complementary to SEQ ID NO:1, as is the nature of double stranded DNA. Nevertheless, solely in order to expedite prosecution of the instant application, Applicants have amended Claim 4 (and thus dependent Claim 5, and its dependent Claim 8) to eliminate reference to Claims 1(d), 2(f), and 3(h) (now Claims 1(d), 2(d), and 3(f)). Applicants therefore contend that Claim 8 as amended is no longer indefinite and respectfully request reconsideration and withdraw of this rejection.

(C) The Office Action has rejected Claim 8, and dependent Claims 9-10, under 35 U.S.C. § 112, second paragraph for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. Specifically, the Office Action asserts that Claim 8 is indefinite in that it only recites the polypeptide of interest by an arbitrary name. Applicants respectfully traverse this assertion.

Applicants note that the Applicant may be his or her own lexicographer, provided that any special meaning assigned to a term is made sufficiently clear in the specification. MPEP 2111.01. In the specification on page 16, lines 12-23, Applicants have defined the term "LGR8 polypeptide" as "a polypeptide comprising the amino acid sequence of any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 and related polypeptides." "Related polypeptides" are similarly defined in the same paragraph. Applicants contend that as the term "LGR8 polypeptide" is explicitly defined in the specification, Claim 8 is not indefinite in reciting this term. Thus Applicants respectfully request that this rejection be withdrawn.

Applicants, believing that the rejection of the pending claims based on 35 U.S.C. §112, second paragraph, for indefiniteness has been overcome by amendment or traversed by argument, respectfully request that this ground of rejection be withdrawn.

4. Claim Rejections – 35 U.S.C. § 101

The Office Action asserts a rejection of claims 1-12 and 43-45 under 35 U.S.C. § 101, stating that the claimed invention is not supported by a specific and substantial asserted utility or a well-established utility. Specifically, the Office Action states that the application does not disclose the biological role of the nucleic acid, the encoded protein, or the significance of either. Applicants respectfully traverse this rejection.

Applicants first note that the Office Action asserts that the instant situation is directly analogous to that which was addressed in *Brenner v. Manson*. Applicants respectfully disagree that the issues of utility of the instant application are directly analogous to those of *Brenner*. In *Brenner*, the Supreme Court determined that claims to a chemical process (and not the chemicals themselves) for producing steroids belonging to a particular class of steroids (said class of steroids comprising one known member previously proven effective in inhibiting tumors in mice) lacked patentable utility because the applicants had not disclosed a sufficient likelihood that the steroids produced by the claimed process had similar tumor-inhibiting properties. As stated in *Brenner*, those applicants disclosed nothing more than (a) a process for producing steroids, and (b) that the compounds produced by the claimed process were homologues of a single known compound shown to have tumor-inhibiting properties. The Court's rationale was that excluding others from making, selling, and, most importantly, using the claimed methods would extend patent protection to the undisclosed and unknown compounds. The instant claims are different in almost every way. First, they are composition of matter claims, so there is no similar global inhibition of technological progress as was present in *Brenner*. Second, the instant application affirmatively teaches specific nucleic acid molecules encoding polypeptides that were found to be actually expressed in animals, primarily in skeletal muscle and in the uterus. Applicants contend, therefore, that the instant application provides the public with a specific benefit (*i.e.*, a particular member of the glycoprotein hormone receptor subfamily). This situation is wholly unlike the circumstances in *Brenner*, where the chemical process of *Brenner* produced a class of compounds which might not have been produced in nature

and which might have had no useful function whatsoever. Under these circumstances, the pending claims do not improperly "engross what may prove to be a broad field." *Brenner*, 383 U.S. at 534-35.

Applicants next note that the Office Action has asserted that Applicants are required to teach that the claimed polypeptides (and nucleic acids encoding the polypeptides) are diagnostic for a specific disease. Applicants respectfully traverse, and contend that this is not their burden in satisfying the requirements of 35 U.S.C. § 101.

Such requirements are set forth in the *Revised Interim Utility Guidelines Training Materials* ("Training Materials"). Specifically, Applicants are required to demonstrate that the asserted utility is specific and substantial, and if so, whether such asserted utility is credible. Applicants contend that they have met this burden. Under the guidelines of the Revised Interim Utility Guidelines Training Materials ("Training Materials"), page 9, Applicants, in the absence of a well-established utility, must first make an assertion of utility for the invention. As the Office Action has recognized (page 6-7 of the Office Action), the Applicants have asserted that the claimed nucleic acid molecules, or the polypeptides they encode, may be used to treat, diagnose, ameliorate, or prevent a number of diseases, disorders, or conditions associated with LGR8 polypeptides, for example using the claimed molecules to diagnose or treat diseases and conditions that modulate cellular proliferation and differentiation, such as wound healing. Thus Applicants have made an assertion of utility for the invention.

Next, the assertion of utility must identify a specific utility. The Training Materials, on page 5, define a "specific utility" as a utility that is *specific* to the subject matter claimed, as contrasted with a *general* utility that would be applicable to the broad class of the invention. To illustrate the difference between a specific utility and a general utility, the Training Materials refer to a claim directed to a polynucleotide whose only asserted utility is that of a gene probe or a chromosome marker, which is a use that all polynucleotide sequences would have, and therefore, is merely a general utility. As applied to the instant application, the claimed subject matter encompasses nucleic acid molecules encoding LGR8 polypeptides, while the broad class of the invention is nucleic acid molecules. The present application asserts a utility that not *all* polynucleotide sequences would have, *i.e.*, not *all* polynucleotide sequences could be used to treat, diagnose, ameliorate, or prevent a number of diseases, disorders, or conditions associated with LGR8 polypeptides. Thus, Applicants

contend that the asserted utility is *specific* to the subject matter claimed, and thus satisfies the first prong of a utility analysis.

Third, the assertion of utility must be substantial. The Training Materials, on page 6, define a "substantial utility" as a utility that has a "real world" use. Members of the glycoprotein hormone receptor family are well known and play an important role in multiple disease states and conditions (see page 83 of the specification, line 25 *et. seq.*). A member of the glycoprotein hormone receptor family has a "real world" use in binding glycoprotein ligands and thus in treating, diagnosing, ameliorating, or preventing various disease states and conditions associated with glycoprotein signaling. Thus, Applicants contend that the asserted utility is substantial, and thus satisfies the second prong of a utility analysis.

Finally, the assertion of utility must be credible. The Training Materials, on page 5, define a "credible utility" as an assertion of utility that is believable to one of ordinary skill in the art based on the totality of evidence and reasoning provided. Furthermore, the Training Materials state that a credible utility is assessed from the standpoint of whether a person of ordinary skill in the art would accept that the recited or disclosed invention is currently available for such use. The instant application teaches nucleotide sequences encoding an amino acid sequence for human and murine LGR8 polypeptides (Figures 1-8). The specification also teaches that the LGR8 polypeptides of the invention bear a high homology to LGR7 (page 91, line 1 and Figure 10). Specifically, the polypeptide encoded by SEQ ID NO:1 is 51.6% identical to LGR 7 (*see* Exhibit A, enclosed herewith). Furthermore, the specification teaches that the LGR8 polypeptides of the invention comprise a large N-terminal leucine-rich repeat-containing extracellular domain, and, with the exception of LGR8-D, a series of 7 transmembrane domains and a cytoplasmic C-terminal region (pages 90-91). One of skill in the art would clearly recognize that the polynucleotides of the instant application encode polypeptides that are homologous members of the LGR family, which share these common structural features (see Hsu *et al.* 1998, submitted with Applicants IDS of June 5, 2002 and enclosed herewith for the convenience of the Examiner as Exhibit B). In fact, the Office Action itself has recognized that the instant protein is a member of the LGR family (page 8). Thus, based on the totality of the evidence, one of ordinary skill in the art would find the asserted utility, *i.e.*, the use to treat, diagnose, ameliorate, or prevent a number of diseases, disorders, or conditions associated with LGR8 polypeptides, to be believable. Furthermore, a person of ordinary skill in the art,

recognizing that the disclosed invention is a member of the LGR family, would accept that it is currently available for use as a glycoprotein hormone receptor. Thus, Applicants contend that the asserted utility is credible to one of ordinary skill in the art, and satisfies the third prong of a utility analysis.

Applicants respectfully submit that because the instant application contains an assertion of a specific and substantial utility for the claimed invention that would be credible to one of skill in the art, the rejection under 35 U.S.C. § 101 should be withdrawn.

2. Claim Rejections – 35 U.S.C. § 112, first paragraph

(A) The Office Action asserts a rejection of claims 1-12 and 43-45 under 35 U.S.C. § 112, first paragraph, as containing subject matter that was not described in such a way as to enable one of skill in the art to which it pertains, or with which it is most clearly connected, to make and use the invention. The Office Action states that because the claimed invention is not supported by a specific and substantial asserted utility or a well-established utility, one skilled in the art would not know how to use the claimed invention.

Applicants contend that this ground of rejection stands or falls with the rejection asserted in the Office Action under 35 U.S.C. § 101. As set forth above, Applicants have provided affirmative evidence that the asserted utility would be credible to one of ordinary skill in the art. Applicants respectfully contend that because the instant application in fact contains an assertion of a specific and substantial utility for the claimed invention that one of ordinary skill in the art would find to be credible, this rejection under 35 U.S.C. §112, first paragraph, is overcome, and should be withdrawn.

(B) The Office Action asserts a rejection of claims 1-3 under 35 U.S.C. § 112, first paragraph, as not being described in the specification in such a way as to reasonably convey to one skilled in the art that the inventors, at the time the application was filed, had possession of the claimed invention. Specifically, the Action states that the claimed nucleic acid molecules, other than SEQ ID NO:1 and the nucleotide sequences encoding the polypeptide of SEQ ID NO: 2, encompass variant sequences which were not originally contemplated. Applicants respectfully traverse this assertion.

Applicants note that the written description requirement for a claimed genus "may be satisfied through sufficient description of a representative number of species by actual reduction to

practice... ***or*** by disclosure of the relevant, identifying characteristics ... sufficient to show the applicant was in possession of the claimed genus." (Emphasis added). MPEP 2163(II)(A)(3)(a)(ii). Applicants contend that the instant application, in combination with the knowledge already possessed by of one of skill in the art, does in fact disclose such relevant, identifying characteristics.

The claims as amended include SEQ ID NO: 1 and fragments thereof, (or polynucleotides encoding SEQ ID NO: 2 and SEQ ID NO: 3 and fragments thereof), polynucleotides that hybridize to these sequences, polynucleotides encoding conservatively substituted variants of SEQ ID NO: 2 and SEQ ID NO: 3, and complementary sequences. The instant application teaches the nucleotide sequence and corresponding amino acid sequence for a human LGR8 polypeptide. The specification specifically discloses SEQ ID NO: 1, SEQ ID NO: 2, and SEQ ID NO: 3, and inherently discloses fragments thereof – as fragments are merely portions of the specifically disclosed full-length sequences. Moreover, the specification positively recites that fragments of the disclosed sequences are encompassed within the scope of the invention. The specification also teaches fragments that hybridize to these sequences and recites the specific hybridization conditions of the amended claims. With respect to complementary sequences, the specification discloses that such sequences are encompassed within the scope of the invention; furthermore, one of skill in the art would readily be able to envision the detailed structure of such sequences.

In addition, the instant application teaches that conservative amino acid substitutions will produce a polypeptide having functional and chemical characteristics similar to those of LGR8 polypeptides (page 25, lines 16-21) and that one of skill in the art will be able to determine suitable variants of SEQ ID NO: 2 or SEQ ID NO: 3 using well-known techniques (page 28). The application further discloses (page 28, line 6 thru page 30, line 25) many of these well-known techniques. One of skill in the art, using the teachings of the specification, would be able to predict a structure for the LGR8 polypeptides of the instant application and to deduce which residues specifically are important for ligand binding. Such studies would be a matter of routine experimentation and would allow one of ordinary skill in the art to prepare LGR8 polypeptides that have functional and chemical characteristics similar to those of SEQ ID NO: 2 or SEQ ID NO: 3.

Applicants, believing that the rejection of the pending claims based on 35 U.S.C. §112, first paragraph, for lack of written description have been overcome by amendment or traversed by argument, respectfully request that this ground of rejection be withdrawn.

(C) The Office Action further asserts a rejection of claims 1-3 under 35 U.S.C. § 112, first paragraph, as not being enabled. Applicants respectfully traverse this assertion.

Claim 1 as amended recites explicitly-disclosed SEQ ID NO: 1, nucleic acid molecules encoding specifically disclosed SEQ ID NOs: 2 and 3, nucleic acid molecules that hybridize to these sequences under specific hybridization conditions, and nucleic acid molecules complementary to these sequences. The Office Action (page 12) has recognized that the specification is enabling for SEQ ID NO: 1 and polynucleotides encoding SEQ ID NO: 2 (and presumably SEQ ID NO: 3). With regard to hybridizing sequences, it would be a matter of routine experimentation for one of skill in the art to determine whether a particular sequence hybridized to SEQ ID NO: 1, or sequences encoding SEQ ID NO: 2 or SEQ ID NO: 3, under the recited conditions. Furthermore, one of skill in the art would readily be able to discern complementary sequences, by basic base pairing. Thus Applicants contend that Claim 1 is fully enabled by the specification.

Claims 2 and 3 recite fragments of explicitly-disclosed SEQ ID NO: 1 (or fragments of polynucleotides encoding explicitly-disclosed SEQ ID NOs: 2 or 3), polynucleotides that hybridize to the complement of these nucleotide sequences under specific hybridization conditions, polynucleotides encoding conservatively substituted variants of SEQ ID NO: 2 or SEQ ID NO: 3, and complementary sequences. Applicants note that Claim 2 as amended no longer recites allelic or splice variants. Applicants contend that the specification is fully enabling for fragments of the explicitly disclosed sequences (including C- and/or N- terminal truncations), as one of skill in the art would, with little effort, be able to create a sub-section of the explicitly disclosed sequences. Furthermore, as discussed above, one of skill in the art would readily be able to discern complementary sequences, or, with routine experimentation, determine whether a particular sequence hybridized to the recited sequences.

With regard to conservative amino acid substitutions, contrary to the assertion of the Office Action, the specification does in fact provide guidance to enable one of ordinary skill in the art to determine, without undue experimentation, the positions in the protein which are tolerant to change. The specification provides one of skill in the art with guidelines for preparing a polypeptide having one or more conservative substitutions (and thus the polynucleotide encoding such a polypeptide) that retains the functional and chemical characteristics of the LGR8 polypeptides. The specification provides exemplary conservative amino acid substitutions (Table 1, pages 27-28), and teaches that

amino acid substitutions may be made within a class of side chains (page 26, line 8), further considering the hydropathic (page 26, line 23) or hydrophilic (page 27, line 12) indices of amino acids. With respect to selection of specific amino acids, one of skill in the art may compare similar sequences within a species or from species to species (page 28, line 6) to identify regions likely to sustain alterations in amino acid sequence while maintaining functional and chemical characteristics. Additionally, one of skill in the art may analyze structure-function studies (page 29, line 14) or three-dimensional structures (page 29, line 21) in predicting which amino acid residues are important for activity or structure. Thus, Applicants contend that the determination of LGR8 variants having conservative substitutions is well within the skill of one of ordinary skill in the art through the practice of nothing more that routine experimentation.

Applicants, believing that the rejection of the pending claims based on 35 U.S.C. §112, first paragraph, for lack of enablement has been overcome by amendment or traversed by argument, respectfully request that this ground of rejection be withdrawn.

CONCLUSIONS

Applicants respectfully contend that all conditions of patentability are met in the pending claims as amended and therefore respectfully request allowance.

If Examiner Seharaseyon believes it to be helpful, Examiner Seharaseyon is invited to contact the undersigned representative by telephone at (312) 913-0001.

Respectfully submitted,
McDonnell Boehnen Hulbert & Berghoff

Dated: 7-26-04

By: Sherri L. Oslick
Sherri L. Oslick, Ph.D.
Reg. No. 52,087

GeneStream align Home Page

align Search Help

---

align Results

Please site: *Pearson, W.R., Wood, T., Zhang, Z., and Miller, W. (1997) Comparision of DNA sequences with protein sequences, Genomics 46: 24-36*

```
>_ LGR8A                                                    754 aa vs.
>_ LGR7                                                    757 aa
scoring matrix: , gap penalties: -12/-2
51.6% identity;          Global alignment score: 2615

               10        20        30        40        50        60
654037 MIVFLVFKHLFSLRLITMFFLLHFIVLINVKDFALTQGSMITPSCQKGYFPCGNLTKCLP
       :    ::                : .::  : :.   :. .  .:. ::::::::.::::
_      MTSGSVF----------------FYILIFGKYFSHGGGQDV--KCSLGYFPCGNITKCLP
                               10        20          30        40

               70        80        90       100               110
654037 RAFHCDGKDDCGNGADEENCGDTSGWATIFGTVHGNANSV-------ALTQECFLKQYPQ
       . .::.: ::::: :::.::::.::.  :    ..  ..       : : ::.. . :
_      QLLHCNGVDDCGNQADEDNCGDNNGWSMQFDKYFASYYKMTSQYPFEAETPECLVGSVPV
             50        60        70        80        90        100

           120       130       140       150       160       170
654037 CCDCKETELECVNGDLKSVPMISNNVTLLSLKKNKIHSLPDKVFIKYTKLKKIFLQHNCI
        : :.  ::.: . .:..:: .:.::: .::. : :..::   : .:  :.:..::.: :
_      QCLCQGLELDCDETNLRAVPSVSSNVTAMSLQWNLIRKLPPDCFKNYHDLQKLYLQNNKI
            110       120       130       140       150       160

           180       190       200       210       220       230
654037 RHISRKAFFGLCNLQILYLNHNCITTLRPGIFKDLHQLTWLILDDNPITRISQRLFTGLN
         ::  :: :: .:  :::.:: :: :.:::.:::.: :::..:: ..:::   : :::
_      TSISIYAFRGLNSLTKLYLSHNRITFLKPGVFEDLHRLEWLIIEDNHLSRISPPTFYGLN
            170       180       190       200       210       220

           240       250       260       270       280       290
654037 SLFFLSMVNNYLEALP-KQMCAQMPQLNWVDLEGNRIKYLTNSTFLSCDSLTVLFLPRNQ
       ::..: ..:: :  :: : .: .::.:.:.:::::.:. : : ::.::..:::: . .:.
_      SLILLVLMNNVLTRLPDKPLCQHMPRLHWLDLEGNHIHNLRNLTFISCSNLTVLVMRKNK
            230       240       250       260       270       280

           300       310       320       330       340       350
654037 IGFVPEKTFSSLKNLGELDLSSNTITELSPHLFKDLKLLQKLNLSSNPLMYLHKNQFESL
       :. . :.::. :..: ::::.:: : .: : .:::::: :..:::: ::.. .. :::. :
_      INHLNENTFAPLQKLDELDLGSNKIENLPPLIFKDLKELSQLNLSYNPIQKIQANQFDYL
            290       300       310       320       330       340

           360       370       380       390       400       410
654037 KQLQSLDLERIEIPNINTRMFQPMKNLSHIYFKNFRYCSYAPHVRICMPLTDGISSFEDL
```

```
          .:.::.:: ::: ::. :::.:. ::::::::::.:.::::::: : : ::::::::.:.:
_         VKLKSLSLEGIEISNIQQRMFRPLMNLSHIYFKKFQYCGYAPHVRSCKPNTDGISSLENL
               350       360       370       380       390       400

               420       430       440       450       460       470
654037    LANNILRIFVWVIAFITCFGNLFVIGMRSFIKAENTTHAMSIKILCCADCLMGVYLFFVG
          ::. : :.::::::. .::::::::: :: .:..:: .:::: ::::::::::::: .:
_         LASIIQRVFVWVVSAVTCFGNIFVICMRPYIRSENKLYAMSIISLCCADCLMGIYLFVIG
               410       420       430       440       450       460

               480       490       500       510       520       530
654037    IFDIKYRGQYQKYALLWMESVQCRLMGFLAMLSTEVSVLLLTYLTLEKFLVIVFPFSNIR
           ::.:.::.:.:.: ::::::::::.:: ::.:::::::::::::::::.:::::.. ::.::  .:
_         GFDLKFRGEYNKHAQLWMESTHCQLVGSLAILSTEVSVLLLTFLTLEKYICIVYPFRCVR
               470       480       490       500       510       520

               540       550       560       570       580       590
654037    PGKRQTSVILICIWMAGFLIAVIPFWNKDYFGNFYGKNGVCFPLYYDQTEDIGSKGYSLG
          ::: .: ..:: ::..::..: ::. ::..: :.:: ::::::::. ..::.::.. ::..
_         PGKCRTITVLILIWITGFIVAFIPLSNKEFFKNYYGTNGVCFPLHSEDTESIGAQIYSVA
               530       540       550       560       570       580

               600       610       620       630       640       650
654037    IFLGVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWI
          ::::.:: ::.::::::: .:: :....:. .::.:: .:. .:.:::::::.::.:::
_         IFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVFTDALCWI
               590       600       610       620       630       640

               660       670       680       690       700        710
654037    PVFVVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKLKQLLHKH-QR
          :.::::::::..:::: :.:::.:::::::::::::::::  ::. .... ... ::
_         PIFVVKFLSLLQVEIPGTITSWVVIFILPINSALNPILYTLTTRPFKEMIHRFWYNYRQR
               650       660       670       680       690       700

                720       730              740             750
654037    KSIFKIKKKSLSTSIVWIE-------DSSSLKLGVLN---KITL--GDSIMKPVS
          ::. .  .:. . :..:.:           .:  ...   ...:   .. ..  :
_         KSMDSKGQKTYAPSFIWVEMWPLQEMPPELMKPDLFTYPCEMSLISQSTRLNSYS
               710       720       730       740       750

Elapsed time:  0:00:00
```

Characterization of Two LGR Genes Homologous to Gonadotropin and Thyrotropin Receptors with Extracellular Leucine-Rich Repeats and a G Protein-Coupled, Seven-Transmembrane Region

Sheau Yu Hsu*, Shan-Guang Liang*, and Aaron J. W. Hsueh

Division of Reproductive Biology
Department of Gynecology and Obstetrics
Stanford University Medical School
Stanford, California 94305-5317

The receptors for LH, FSH, and TSH belong to the large G protein-coupled, seven-transmembrane (TM) protein family and are unique in having a large N-terminal extracellular (ecto-) domain containing leucine-rich repeats important for interaction with the glycoprotein ligands. We have identified two new leucine-rich repeat-containing, G protein-coupled receptors and named them as LGR4 and LGR5, respectively. The ectodomains of both receptors contain 17 leucine-rich repeats together with N- and C-terminal flanking cysteine-rich sequences, compared with 9 repeats found in known glycoprotein hormone receptors. The leucine-rich repeats in LGR4 and LGR5 are arrays of 24 amino acids showing similarity to repeats found in the acid labile subunit of the insulin-like growth factor (IGF)/IGF binding protein complexes as well as slit, decorin, and Toll proteins. The TM region and the junction between ectodomain and TM 1 are highly conserved in LGR4, LGR5, and seven other LGRs from sea anemone, fly, nematode, mollusk, and mammal, suggesting their common evolutionary origin. In contrast to the restricted tissue expression of gonadotropin and TSH receptors in gonads and thyroid, respectively, LGR4 is expressed in diverse tissues including ovary, testis, adrenal, placenta, thymus, spinal cord, and thyroid, whereas LGR5 is found in muscle, placenta, spinal cord, and brain. Hybridization analysis of genomic DNA indicated that LGR4 and LGR5 genes are conserved in mammals. Comparison of overall amino acid sequences indicated that LGR4 and LGR5 are closely related to each other but diverge, during evolution, from the homologous receptor found in snail and the mammalian glycoprotein hormone receptors. The identification and characterization of new members of the LGR subfamily of receptor genes not only allow future isolation of their ligands and understanding of their physiological roles but also reveal the evolutionary relationship of G protein-coupled receptors with leucine-rich repeats. (Molecular Endocrinology 12: 1830–1845, 1998)

0888-8809/98/$3.00/0
Molecular Endocrinology

INTRODUCTION

Proteins in the large seven-transmembrane (TM), G protein-coupled receptor (GPCR) superfamily are functionally diverse and include receptors ranging from the cAMP receptor in slime mold to mammalian neurotransmitter and glycoprotein hormone receptors (1–5). Agonist occupancy of these plasma membrane proteins leads to the activation of different G proteins, which in turn modulate the activity of different effector enzymes and ion channels (6, 7). Gonadotropins (LH, FSH, CG) and TSH are essential for the growth and differentiation of the gonads and thyroid gland, respectively. These glycoprotein hormones bind specific membrane-bound GPCRs on target cells to activate the Gs-cAMP-protein kinase A pathway (8–11). The glycoprotein hormone receptors represent a subgroup of GPCRs that have a large N-terminal extracellular (ecto-) domain containing leucine-rich repeats important for interaction with glycoprotein hormones from adenohypophysis and placenta, which leads to cAMP production in target cells.

Based on the conserved sequences of putative glycoprotein hormone receptors in *Drosophila* and sea anemone (12, 13), the expression sequence tags (EST)

in the GenBank were searched, and fragments of two new mammalian receptors in this subfamily of leucine-rich repeat-containing, G-protein-coupled receptor (LGR) were identified.[1] We report here the molecular cloning of these putative mammalian receptors with a protein architecture that is similar to the known glycoprotein hormone receptors and their invertebrate homologs in both ectodomains and TM segments. In addition to the three known receptors, the ectodomains of LGR4 and LGR5 show high homology with the acid labile subunit (ALS) (14–16), slit (17), decorin (18), and Toll proteins (19) containing leucine-rich repeats, suggesting a common evolutionary origin. In contrast to the restricted tissue expression pattern of known gonadotropin and TSH receptors, these new receptors were found in multiple tissues. Identification of this expanding family of LGRs has implications for future studies to identify putative ligands for these orphan receptors and for the understanding of the evolutionary origin of proteins in this expanding subfamily of leucine-rich repeat-containing seven-TM receptors.

RESULTS

Conserved Architecture of Ectodomain, TM Region, and C-Terminal Tail of LGR4 and LGR5

Human sequences related to the sea anemone and *Drosophila* glycoprotein hormone receptors (12, 13) were identified from the EST database based on their similarities to receptors found in the lower species and nonidentity to the three known human glycoprotein hormone receptors. The full-length cDNAs for novel receptors were isolated using RT-PCR and repeated screening of sublibraries from rat ovary or human placenta enriched with each receptor cDNA. Positive clones with long inserts were sequenced and aligned to identify the open reading frames (ORFs) of individual receptors. The prototypic LGR consists of an ectodomain with leucine-rich repeats and a C-terminal half with seven-TM domains similar to other GPCRs. Because three known glycoprotein hormone receptors have the same leucine-rich repeat-containing ectodomain and G protein-coupled TM region, the new mammalian receptors were named LGR4 and LGR5, respectively.

LGR4 cDNA from rat ovary consists of 3,504 bp with a predicted ORF of 951 amino acids, whereas LGR5 from human placenta has 4,208 bp with a 907-amino acid ORF (Fig. 1). The ectodomains of LGR4 and LGR5 are more closely related to each other (54% identity; 67% similarity) than to the three known LGRs (18–23% identity; 33–35% similarity). Similar to three known glycoprotein hormone receptors, LGR4 and LGR5 are characterized by multiple leucine-rich repeat sequences (Fig. 1C, Table 1, and Ref. 18). Six and four consensus N-linked glycosylation sites (Fig. 1C, *underlined N*; and Table 1) were found in the ectodomains of LGR4 and LGR5, respectively, and two of these sites were conserved between LGR4 and LGR5.

Although there are 17 leucine-rich repeats in LGR4 and LGR5 as compared with 9 repeats in the glycoprotein hormone receptors, alignment of the N-terminal 9 repeats of the five mammalian LGRs showed that the third potential N-glycosylation site in LGR4 (Asn 199) and the second potential N-glycosylation site in LGR5 (Asn 208) align perfectly with the N-glycosylation site found in the sixth repeat of gonadotropin and TSH receptors. In addition, clusters of cysteines (cysteine-rich sequences) are present in the N-terminal region and the junction between the ectodomain and TM 1. Because the four N-flanking cysteine residues are conserved in all of the mammalian LGRs and other leucine-rich repeat proteins (Fig. 1B), these residues likely form disulfide bonds essential for maintaining the conformation of the large ectodomain of these receptors. In the C-flanking region, LGR4 and LGR5 also contain a cysteine-rich, chemokine-like region similar to the consensus CF3 subtype domain recently identified in 45 glycoprotein hormone receptors isolated from different mammals (20, 21), further confirming the similar protein architecture of these receptors. In particular, the core sequences of this consensus CF3 domain (CCAF and FK/NPCE sequences) are completely conserved (Fig. 1D). However, the length of residues between conserved cysteines in LGR4 and LGR5 (CC-4X-C-4/54X-C) in this region is different from that found in the three known glycoprotein hormone receptors (CC-15/23X-C-31/88X-C). In addition, the junctional insertion of about 50 amino acids unique for the TSH receptor (22) was missing in LGR4 and LGR5.

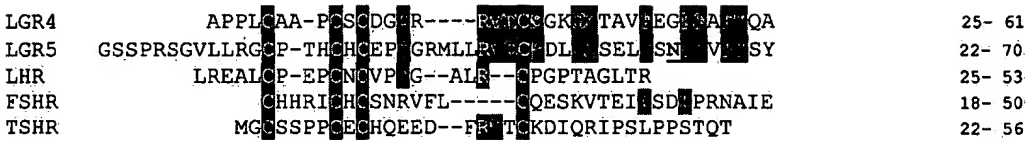
Seven membrane-spanning regions were predicted based on stretches of hydrophobic amino acids forming α-helices (SOUSI server, www.tuat-.ac.jp/cgi/~mitaku/& NAKAI server, http://psort.nibb.ac.jp/cgi-bin; Fig. 1E). They are believed to delimit a barrel-like cylinder structure with the apolar face of the helices turned toward the membrane lipids. Similar to their ectodomains, the TM helices of LGR4 and LGR5 are more homologous to each other (49% identity; 64% similarity) than to the known LGRs (25–27% identity; 48–52% similarity). In contrast to the TM helices, the sequences in intracellular loop 3, an area believed to be important for G protein coupling for adrenergic receptors (23), are similar between the two new receptors (54% identity; 73% similarity) but distinct from the three known glycoprotein hormone receptors (18% identity; 36% similarity). Likewise, three outside and two other intracellular loops of LGR4 and LGR5 show closer homology to each other as compared with gonadotropin and TSH receptors. The highly con-

[1] GenBank accession numbers for LGR4 and LGR5 are AF061443 and AF061444, respectively.

A. Signal peptide

Rat	LGR4	MPGPLGLLCFLALGLLGSAGPSGA	1- 24
Human	LGR5	MDTSRLGVLLSLPVLLQLATG	1- 21
Human	LHR	MKQRFSALQLLKLLLLLQPPLPRA	1- 24
Human	FSHR	MALLLVSLLAFLSLGSG	1- 17
Human	TSHR	MRPADLLQLVLLLDLPRDLGG	1- 21

B. N-flanking cysteine-rich sequence



C. Leucine-rich repeats

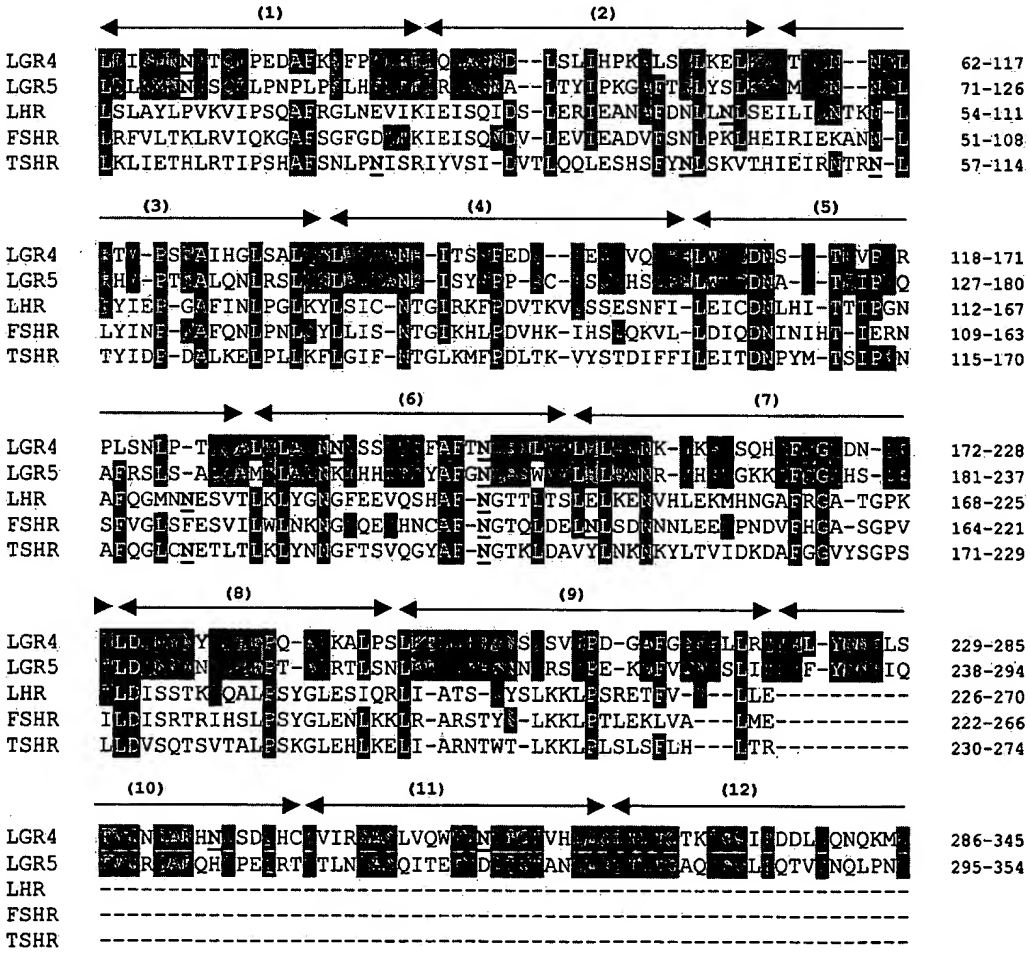
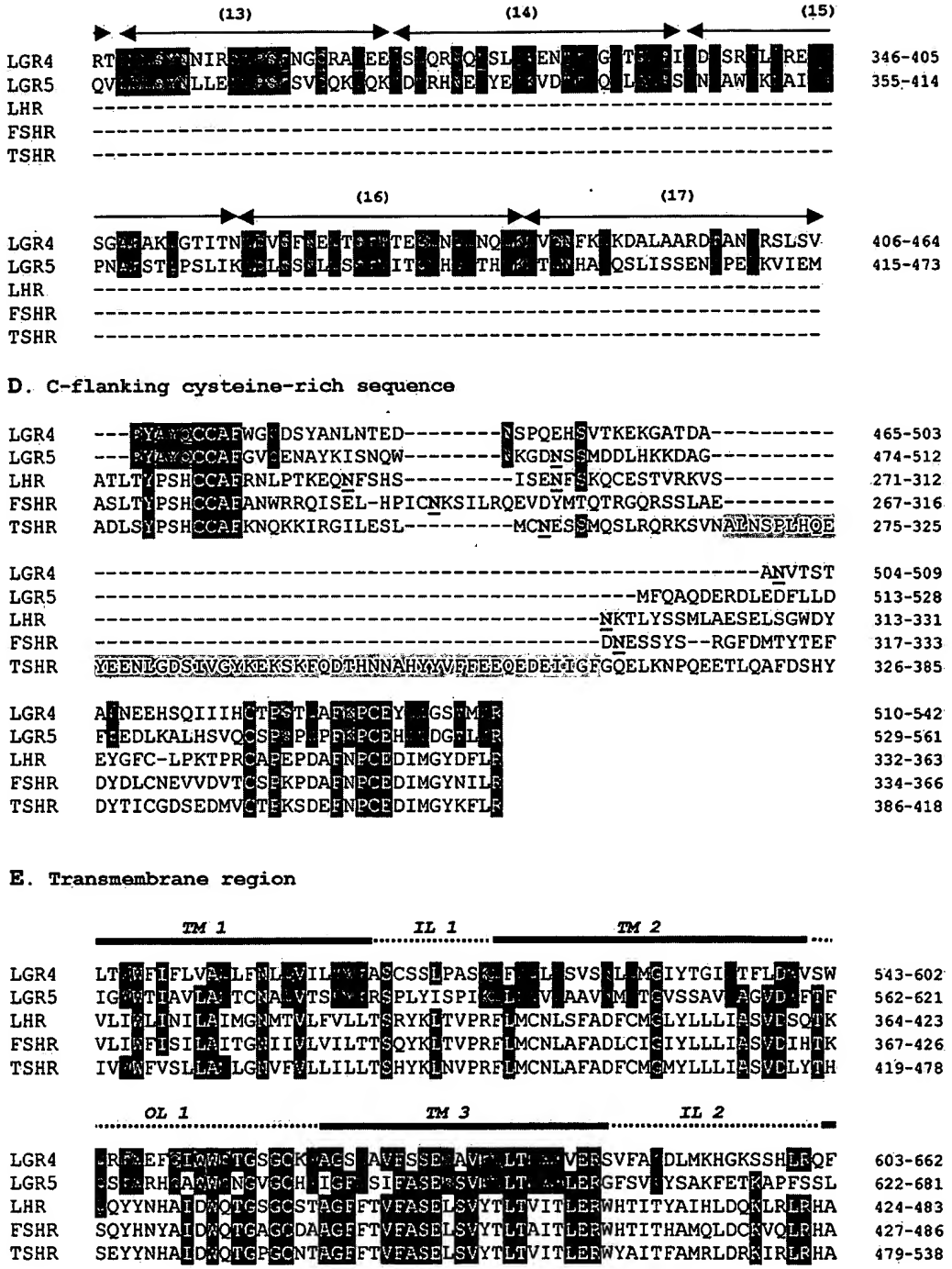
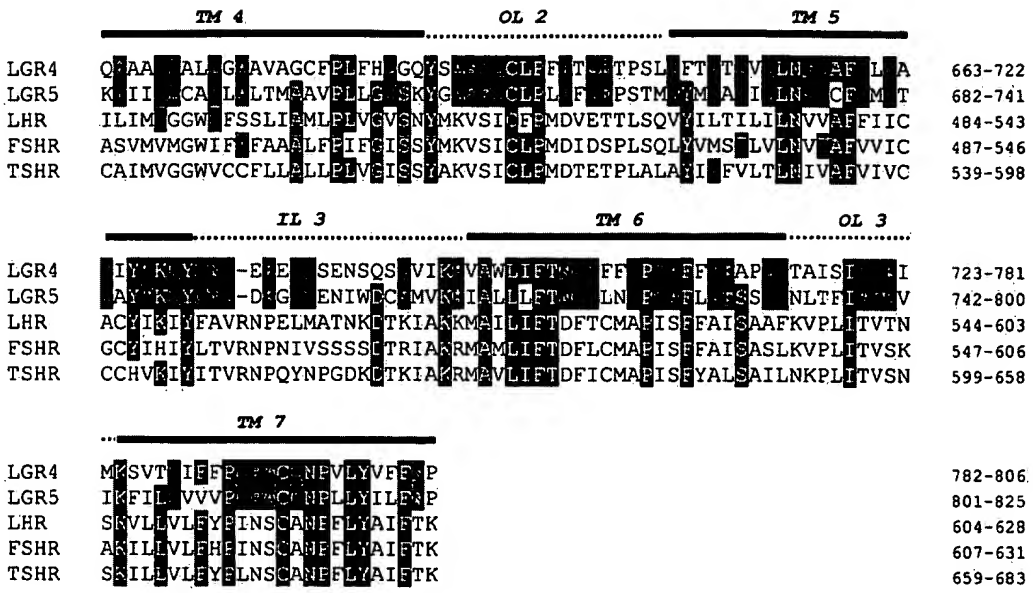


Fig. 1. Comparison of Deduced Amino Acid Sequence of LGR4 and LGR5 cDNAs and Those Encoding FSH, LH, and TSH Receptors and Other Leucine-Rich Repeat-Containing Proteins

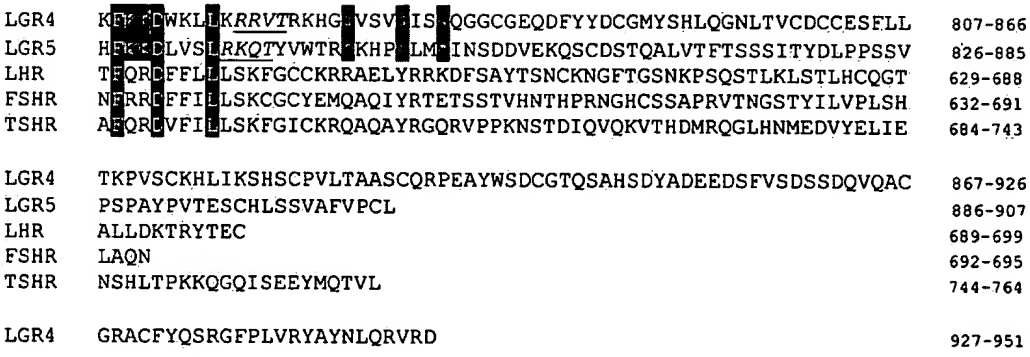
Sequence alignment for different regions of LGR4, LGR5, and three human glycoprotein hormone receptors were performed. *Heavily shaded* residues are identical in at least four of the five receptor proteins shown, whereas the lightly shaded residues show identity in LGR4 and LGR5. Amino acid numbers are shown on the *right*. Asparagine residues in the potential N-linked glycosylation sites are *underlined and shown in bold*. Gaps introduced for optimal alignment are indicated as *dashes*. A, Signal peptide. B, N-flanking cysteine-rich sequence. C, Leucine-rich repeats. Seventeen leucine-rich repeat sequences in LGR4 and LGR5 are indicated by *arrows* at the *top* of the sequences. Because there are only nine repeat sequences in the three glycoprotein hormone receptors, they were arbitrarily aligned with the first nine repeats in LGR4 and LGR5. D, C-flanking cysteine-rich sequence. E, TM region. Seven putative TM (TM) domains are indicated at the *top* of the sequence. Intracellular loops (IL) and outside loops (OL) are also indicated. F, C-terminal tail. Conserved protein kinase A phosphorylation sites are *underlined*. The



GenBank accession numbers for LGR4 and LGR5 are AF061443 and AF061444, respectively. G, Comparison of leucine-rich repeats found in LGRs and diverse other proteins with typical type leucine-rich repeats. Consensus leucine-rich repeats in LGR4 and LGR5 were compared with those found in three human glycoprotein hormone receptors (LHR/FSHR/TSHR) and LGRs from lower species. In addition, the leucine-rich repeats of several other secretory proteins (ALS of IGF/IGF binding protein complexes, slit, and decorin) and single-TM domain receptors (Toll and Tartan) (36, 62) with homologous repeats are shown. The number of repeats found in each protein is shown in *parenthesis*, whereas consensus amino acid residues are *shaded*. *Uppercase letters* indicate more than half of the leucine-rich repeats found in a given protein are conserved, whereas *lowercase letters* denote less than 50% consensus. a, Aliphatic residues, c, charged residues; r, rat; h, human; Ae, *Anthopleura elegantissima* (sea anemone); Ce, *Caenorhabditis elegans* (nematode); Dm, *Drosophila melanogaster* (fly); Ls: *Lymnaea stagnalis* (snail).



F. C-terminal tail



G. consensus sequence of LRR proteins

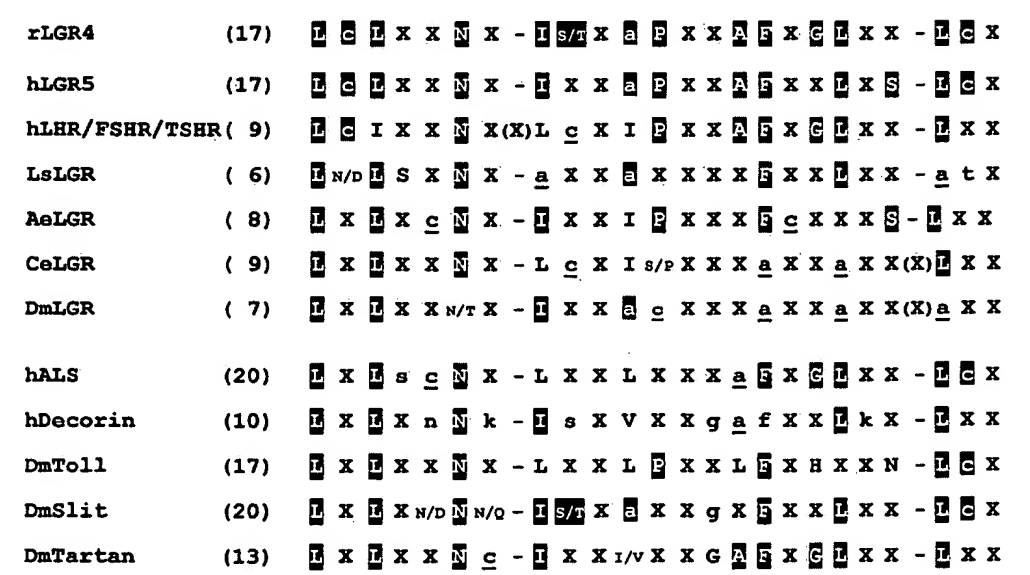


Fig. 1. *continued*

Table 1. Structural Features of LGR4 and LGR5 in Comparison with Human Gonadotropin and TSH Receptors

	Length of Receptor (SP/EC/TM/CT)	No. of Leucine-Rich Repeats	No. of Glycosylation Sites	Chromosomal Localization	Tissue Distribution
FSH receptor	695 (17/349/264/65)	9	4	2p21	Ovary, testis
LH receptor	699 (24/339/264/72)	9	6	2p21	Ovary, testis
TSH receptor	764 (21/397/264/82)	9	6	14q31	Thyroid
LGR4	951 (24/518/263/146)	17	6	5q34-35.1	Ovary, testis, adrenal, spinal cord, thyroid
LGR5	907 (21/540/263/83)	17	4	12q15	Muscle, placenta, spinal cord

SP, Signal peptide; EC, ectodomain; TM, TM domain; CT, C-terminal tail. Sequences of human receptors for FSH, LH, and TSH are based on Refs. 63, 64, and 22, respectively.

served cysteine residues in the first and second outside loops, predicted to form an intramolecular disulfide bridge to constrain protein conformation, were conserved in all LGRs (24, 25). In addition, proline residues in the fourth, sixth, and seventh TM segments, believed to be necessary for proper insertion of the receptor proteins into the membrane (26), were also conserved. Among the outside loops of these receptors, the highest homology was found in the second loop exhibiting a unique β-strand structure.

Although minimal conservation could be found for the five receptors in the C-terminal tail (Fig. 1F), multiple potential phosphorylation sites were found in LGR4 and LGR5 as in glycoprotein hormone receptors. For the two new receptors, a consensus protein kinase A phosphorylation site was conserved (Fig. 1F, *underlined and italic letters*), suggesting possible regulation through cAMP-regulated phosphorylation (27). In LGR5, potential SH2 and SH3 interacting sequences (amino acids 878–881 and 888–891, respectively) were also found (28).

Comparison of Leucine-Rich Repeats in LGR4 and LGR5 with Similar Repeats in Glycoprotein Hormone Receptors and Other Leucine-Rich Repeat-Containing Proteins

The ectodomains of LGR4 and LGR5 are composed of 17 imperfect leucine-rich repeat motifs of 22–24 amino acids in length (Fig. 1, A and G). The new consensus repeat sequences derived from LGR4 and LGR5 are similar to each other with the exception that glycine 18 is more common in LGR4, whereas serine 21 is more common in LGR5 (Fig. 1G). In addition, repeats 10, 11, 12, and 17 in both receptors are distinct from the remaining repeats and show greater deviation from the consensus leucine-rich repeat sequence (18). Of interest, leucine-rich repeats found in LGR4 and LGR5 are closely related to comparable repeats in the three glycoprotein hormone receptors and LGRs from lower species (Fig. 1G). These repeats are also present in ALS of the insulin-like growth factor (IGF)/IGF binding protein complexes, the proteoglycan decorin, the *Drosophila* and mammalian Toll receptors, the *Drosophila*-secreted protein slit, and the *Drosophila* Tartan receptor (Fig. 1G). A consensus asparagine in residue 6 is present in the repeats of all these proteins, a feature unique to the typical type repeats (18). These findings suggest a close evolutionary origin of the leucine-rich repeats in these proteins of diverse structural arrangement and function (20, 21).

Tissue Expression Pattern of LGR4 and LGR5

Northern blot hybridization was performed to analyze the expression pattern of LGR4 and LGR5 mRNAs in diverse human tissues. As shown in Fig. 2A, a major transcript of 5.5 kb for LGR4 is expressed in multiple steroidogenic tissues (placenta, ovary, testis, and adrenal). The mRNA for this putative receptor is also found in spinal cord, thyroid, stomach, trachea, heart, pancreas, kidney, prostate, and spleen. In contrast, the expression pattern of LGR5 mRNA is more restricted (Fig. 2B). A transcript of 4.3 kb, together with a minor transcript of 2.4 kb for LGR5 mRNA, was found to be highest in the skeletal muscle. This transcript is also present in placenta, spinal cord, brain, adrenal, colon, stomach, and bone marrow.

Lack of Gs Stimulation Mediated by Chimeric Receptors Comprising the Ectodomain of LH Receptor and the TM Region of LGR4 or LGR5

Because chimeric receptors among the known glycoprotein hormone receptors have been successfully used to study signal transduction by these proteins (29), cDNAs for chimeric receptors were generated by fusing the ectodomain of human LH receptor with the TM region and C-terminal tail of either LGR4 or LGR5 [named as L(EC)LGR4(TM) and L(EC)LGR5(TM), respectively]. Cells transfected with the plasmid encoding L(EC)LGR4(TM) showed moderate binding to labeled human (h)CG with a dissociation constant (K_d) value similar to that of the wild-type LH receptor

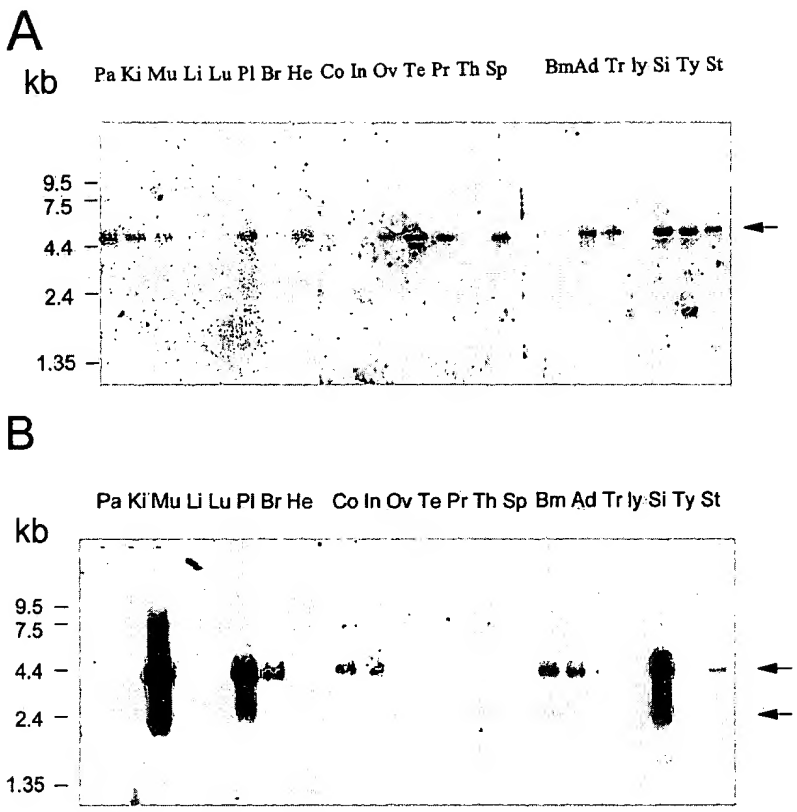


Fig. 2. Expression Pattern of LGR4 and LGR5 mRNA Transcripts in Different Tissues

For Northern blot analysis, 2 μg of poly $(A)^+$-selected RNA from different human tissues were probed with a ^{32}P-labeled LGR4 or LGR5 cDNA probe at 60 C. After washing under stringent conditions, the blots were exposed to x-ray films with an intensifying screen at −80 C for 7 days. Subsequent hybridization with a β-actin cDNA probe was performed to estimate nucleic acid loading (8 h exposure; data not shown). A, LGR4 Northern blot. B, LGR5 Northern blot. Specific LGR transcripts are indicated by *arrows*. Pa, Pancreas; Ki, kidney; Mu, skeletal muscle; Li, liver; Lu, lung; Pl, placenta; Br, brain, He, heart; Co, colon; In, small intestine; Ov, ovary; Te, testis; Pr, prostate; Th, thymus; Sp, spleen; Bm, bone marrow; Ad, adrenal; Tr, trachea; ly, lymph node; Si, spinal cord; Ty, thyroid; St, stomach.

whereas cells transfected with the plasmid encoding L(EC)LGR5(TM) showed lower binding but with high affinity. The B_{max} (ng hCG bound/10^5 cells) and K_d (pM) values for different receptors are: LH receptor, 4.6±3.3 and 195±99; L(EC)LGR4(TM), 2.5±0.8 and 183±114 and L(EC)LGR5(TM), 0.46±0.28 and 549±206, respectively. Despite detectable hCG binding, treatment with increasing doses of hCG did not increase cAMP production by either one of the chimeric receptors. At 10 μg/ml hCG, a 62-fold increase of cAMP production was mediated by the wild type LH receptor but no stimulation of cAMP was found in cells expressing either chimeric receptors ($P < 0.05$).

Isolation of LGR4 and LGR5 Genes, Their Conservation in Vertebrates, and Chromosomal Localization in Humans

Using LGR4 and LGR5 cDNA fragments as probes, a bacterial artificial chromosome-based human genomic DNA library was screened and several genomic clones for LGR4 and LGR5 were isolated. To assess the conservation of the LGR4 and LGR5 genes in diverse vertebrates, Southern blot hybridization of genomic DNA from different species was performed. Under medium stringency washing conditions, the rat LGR4 cDNA and human LGR5 cDNAs hybridized with genomic DNA from all mammalian species tested, suggesting that both LGR4 and LGR5 genes are conserved during mammalian evolution (Fig. 3A).

Genomic fragments (>100 kb) of LGR4 and LGR5 were used as probes in fluorescence *in situ* hybridization (FISH) analysis to identify the chromosomal localization of LGR 4 and LGR5 genes. As shown in Fig. 3B, LGR4 and LGR5 genes were localized to banded DNA in chromosomal 5q34–35.1 and 12q15, respectively.

Conservation of TM and Flanking Regions in Nine LGRs from Diverse Species and the Phylogenetic Relationship of These Receptors

In addition to the three glycoprotein hormone receptors and the two new LGRs discussed here, four similar receptors have been found in lower species. Sequence analysis of mammalian LGRs and homologous receptors from sea anemone (13), fly (12), nem-

atode (30), and snail (31) indicated that the TM region and the junction between ectodomain and TM 1, shown to be important for signal transduction of the known glycoprotein hormone receptors, can be aligned based on BLOCK search (32). In Fig. 4A, BLOCK Maker analysis showed that the TM region and sequences 5′ to TM 1 are highly conserved in all nine receptors and four ungapped blocks can be identified. In addition, the chemical property of residues in this region is highly similar (*lightly hatched* in Fig. 4A). Consensus secondary structure analysis of these receptors further indicated that, in addition to the seven α-helical membrane-spanning domains, one unique β-strand structure could be found in the outside loop 2. Of interest, this region has been shown to be important for the modulation of hormone binding of LH receptor (33); conservation of the secondary structure in this region suggests the outside loop 2 may have a similar role in LGR4 and LGR5.

Further analysis of the phylogenetic relatedness of nine LGRs from diverse species, based on either the full-length receptor sequence (Fig. 4B) or the TM regions (data not shown), suggested that LGR4 and LGR5 diverged early during evolution from the known glycoprotein hormone receptors and from a homologous LGR found in the central nervous system of snail (*Lymnaea stagnalis*; Fig. 4B). In contrast, the three mammalian glycoprotein hormone receptors are more related to the receptor identified in *Drosophila*, and all four of these receptors can be categorized into the same branch of the evolutionary tree together with LGRs found in sea anemone and nematode.

DISCUSSION

The mammalian LGR family of proteins comprises at least five gene products: LH and FSH receptors essential for gonadal development; TSH receptor for thyroid differentiation; and the two new orphan LGR4 and LGR5 without known ligands (Table 1). In contrast to known glycoprotein hormone receptors with 9 leucine-rich repeats in their ectodomain, LGR4 and LGR5 have 17 leucine-rich repeats. Although the TM helices are highly conserved among the five mammalian LGRs, the inside loops of LGR4 and LGR5 are diverged from the glycoprotein hormone receptors. The two new receptors, as FSH, LH, and TSH receptors, are likely to be glycoproteins. In contrast to LH and FSH receptors found in the same region of human chromosome 2p21 (34), LGR4 and LGR5 were localized to distinct human chromosomes. Although the physiological roles of LGR4 and LGR5 are presently unclear, these putative receptors appear to have a wider tissue distribution than gonadotropin and TSH receptors. The availability of LGR4 and LGR5 cDNAs allows future identification of ligands for these orphan receptors and elucidation of their physiological function.

Diverse proteins containing leucine-rich repeats have been identified in prokaryotes, plants, yeast, and many metazoans (18). Leucine-rich repeats represent amphipathic sequences with leucine as the predominant hydrophobic residue and are important for protein-protein interaction (35). The packing of similar repeats allows the formation of a specific hydrogen bond network between neighboring repeats to form a unique secondary structure (21). The leucine-rich repeats in LGR4 and LGR5 belong to the typical type repeats with a conserved asparagine in the middle (18). Conserved cysteine residues flanking leucine-rich repeats are also present in LGR4 and LGR5. Except for Toll-like receptors and a related 18-wheeler receptor (19, 36, 37) containing only a C-terminal cysteine-rich domain, other leucine-rich repeat proteins, like LGRs, have conserved cysteines at both N- and C-flanking regions. These cysteine residues are likely to form disulfide bridges to maintain the overall folding of repeat modules regardless of the number of repeat (21).

Several models for leucine-rich repeats in the ectodomains of mammalian glycoprotein hormone receptors have been postulated (38, 39). These models are based on the crystal structure of the porcine ribonuclease-ribonuclease inhibitor complex in which the repeats of 28 or 29 residues each have an inwardly directed β-sheet (at the concave surface) that might interact with specific ligands and an outwardly directed α-helix (at the convex surface of the horseshoe). The consensus repeat sequences found in LGR4 and LGR5 are most similar to the leucine-rich repeats found in the ALS in the IGF/IGF binding protein complexes important for maintaining the serum IGF reserve (40). They are also similar to the *Drosophila* slit secreted by glia cells in developing neurons (17) and the *Drosophila* and mammalian Toll-like receptors important for dorsal-ventral polarization during embryogenesis and the innate immune responses in adults (19, 36). In addition, a small dermatan sulfate proteoglycan decorin has homologous repeats; this proteoglycan interacts with extracellular matrix and may serve as a reservoir of transforming growth factor-β (TGFβ) (41). All repeats are believed to be involved in protein-protein interactions: RNase inhibitor binds RNase; ALS interacts with IGF-binding protein 3; slit binds laminin; Toll receptor binds Spatzel (42); proteoglycan decorin binds TGFβ and collagen (41, 43); and biglycan binds laminin and fibronectin. For FSH, LH, and TSH receptors, the repeat-containing ectodomains are responsible for binding of cystine-knot fold glycoprotein hormones (38). Based on structural homology with other LGRs, leucine-rich repeats in the ectodomains of LGR4 and LGR5 might also bind specific ligands. Although the putative ligands could be related to known glycoprotein hormones, they could also be related to *Drosophila* Spatzel protein based on the similarity between leucine-rich repeats found in LGR4, LGR5, and the Toll receptors. Of interest, both Spatzel and 8a related ligand Trunk have a conserved cysteine-knot tertiary structure similar to FSH, LH, and TSH

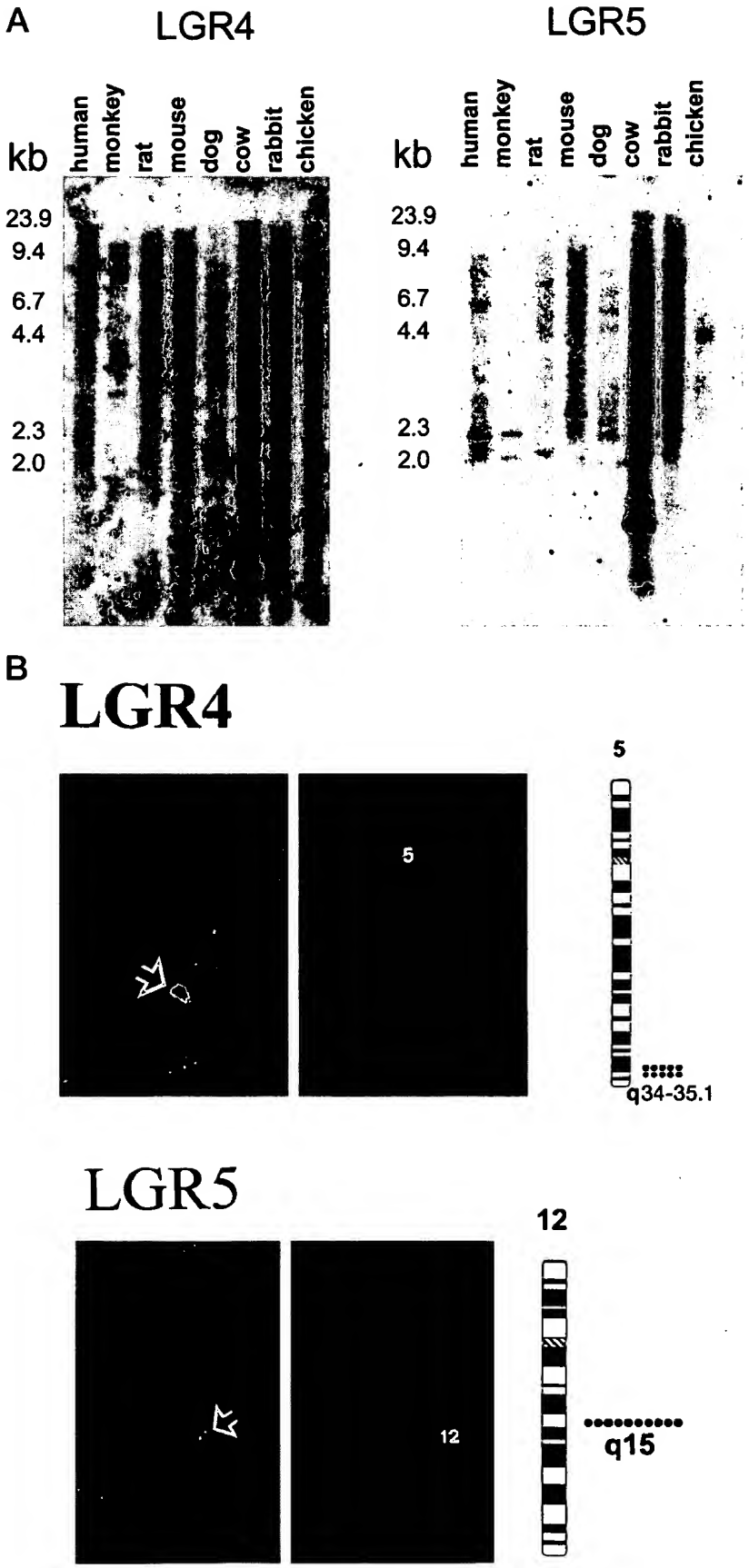


Fig. 3. Conservation of LGR4 and LGR5 Genes in Diverse Vertebrate Species and Their Chromosomal Localization in Human Cells

A, Southern blot hybridization of genomic DNA isolated from different vertebrate species was performed using LGR4 and LGR5

(44). Because leucine-rich repeats in the two novel LGRs are also similar to that found in the ALS of the IGF/IGF binding protein complexes and the proteoglycan decorin, they might also interact with proteins related to IGF-binding protein 3 or TGFβ, ligands for ALS and decorin, respectively.

Both LGR4 and LGR5 contain multiple consensus N-linked glycosylation sites in their ectodomain. In all mammalian LGRs, an N-glycosylation site in leucine-rich repeat 6 was conserved. In addition, the Ala-Phe residues 5′ to this site were also found in LGR4 and LGR5 with the exception of an amino acid insertion. Interestingly, mutation of the conserved Ala to Val in the FSH receptor gene leads to ovarian dysgenesis (45) and spermatogenic failure (46). The conservation of this motif among different LGRs underlines its functional importance.

Alignment of four blocks of homology domains in the TM and flanking regions of nine LGRs from diverse species indicated that multiple α-helices are important for membrane orientation and functional integrity. The first homologous block not only contains TM helix 1 but also extends into ~15 residues in the junction of the ectodomain and the TM region. This conserved region represents a cysteine-rich, chemokine-like structure likely important for correct orientation of the ectodomain to the TM region (47). Based on mutagenesis and chimeric receptor studies, this junction is important for signal transduction and folding of the LH receptor (24, 47, 48). Also, several residues (LGR4 residues: 783K, 791P, and 801Y) at the border and inside the TM helix 7, identified as essential for signal transduction of LH and TSH receptors based on extensive site-directed mutagenesis (8, 47), are highly conserved in all nine LGRs. Sequence alignment further indicated that several key features that distinguish the known glycoprotein hormone receptors from other GPCRs are conserved in the two new LGRs, including the lack of a conserved proline in TM 5, an extra proline in TM 7, and substitution of aromatic residues in TM 5 and 6 of nonglycoprotein hormone receptors with polar residues in LGRs (49–51). However, a third proline residue in TM 7 was found only in LGR4 and LGR5. These data suggest that the seven TM bundles of LGR4 and LGR5 could have similar but distinct spatial orientation as compared with the known glycoprotein hormone receptors. Thus, structural comparison of the expanding group of LGRs could predict the functional importance of critical residues for proper topology of these proteins.

During the preparation of this manuscript, an orphan GPCR (HG38) was reported showing sequence identity to LGR5 except for two amino acids in the ectodomain (52). Using radiation hybrid mapping, HG38 was localized to human chromosome 12q22–23 instead of 12q15 as was found based on the FISH method (Fig. 3B). Although the former method gives greater resolution, further characterization using physical mapping could provide the precise location of LGR5/HG38.

Using a chimeric receptor approach, the signal transduction property of LGR4 and LGR5 was tested. Cells expressing chimeric receptors showed high-affinity hCG binding, but no cAMP stimulation by hCG was detected. Although these findings suggest the two new receptors might not be coupled to the Gs protein, one cannot rule out the possibility that the ectodomain of LH receptor might not be compatible with the exoloops and TM helices of LGR4 and LGR5 for signal transduction. Of interest, a conserved Glu-Arg-Trp triplet motif found in the junction between TM 3 and inside loop 2, postulated to be involved in the interaction between receptors and G proteins (53), is present in the LGR4 and LGR5 but shows substitution in the last residue. In addition, unique SH2 and SH3 interacting sequences, believed to be important for protein-protein interaction in the mitogen kinase cascade (28), were found in the C-terminal tail of LGR5 but not in glycoprotein hormone receptors. The exact ligand-signaling mechanisms for the new LGRs remain to be elucidated.

LGRs most likely represent the evolution of composite proteins or chimeras derived from the duplication of different functional motifs to form protein modules followed by gene rearrangement or exon shuffling (Fig. 5) (54, 55). The basic modules for leucine-rich repeats are stretches of 24 amino acids, whereas the seven-TM region is composed of membrane-spanning α-helical motifs of largely hydrophobic residues. An ancestral gene with leucine-rich repeats could evolve into genes with different functions through gene rearrangement. *Drosophila* slit represents a fusion of leucine-rich repeat domains with an epidermal growth factor domain, whereas genes of the Toll family are derived from the fusion of leucine-rich repeats to the interleukin-1 receptor-like motif (19, 36). The LGR family of proteins represents the fusion of the leucine-rich repeats with an ancestral GPCR. Although closely related to different LGRs, the GRL101 gene found in the central nervous system of snail is unique and may represent a fusion of low-density lipoprotein-binding

cDNA probes. Four micrograms of genomic DNA was digested with the *Eco*RI restriction enzyme and probed with LGR4 or LGR5 cDNA. After hybridization at 60 C, the membrane was washed under medium stringency conditions (0.5% SDS, 0.2 × SSC at 60C) before exposure. B, Using DNA fragments of bacterial artificial chromosome containing human LGR4 and LGR5 genes as probes, chromosomal localization of LGR4 and LGR5 was detected using the FISH method to chromosome 5q34–35.1 and 12q15, respectively. Denatured chromosomes from synchronous cultures of human lymphocytes were hybridized with biotinylated probes for localization. Assignment of the FISH mapping data (*left*) was achieved by superimposing signals with 4,6-diamdino-2-phenylindole-banded chromosome (*center*). Analyses are summarized in the form of human chromosome ideograms (*right*). *Upper panel*, LGR4; *lower panel*, LGR5.

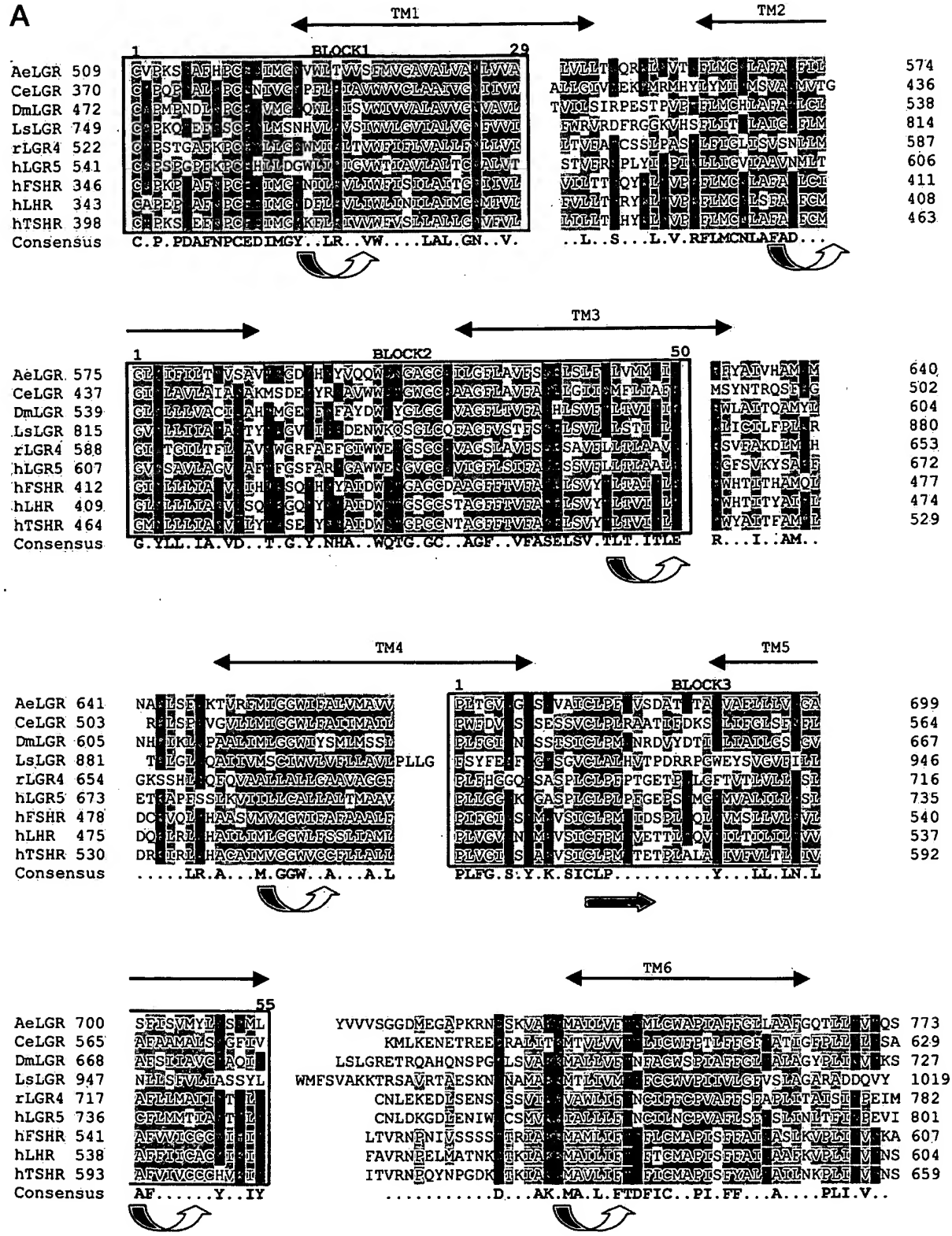
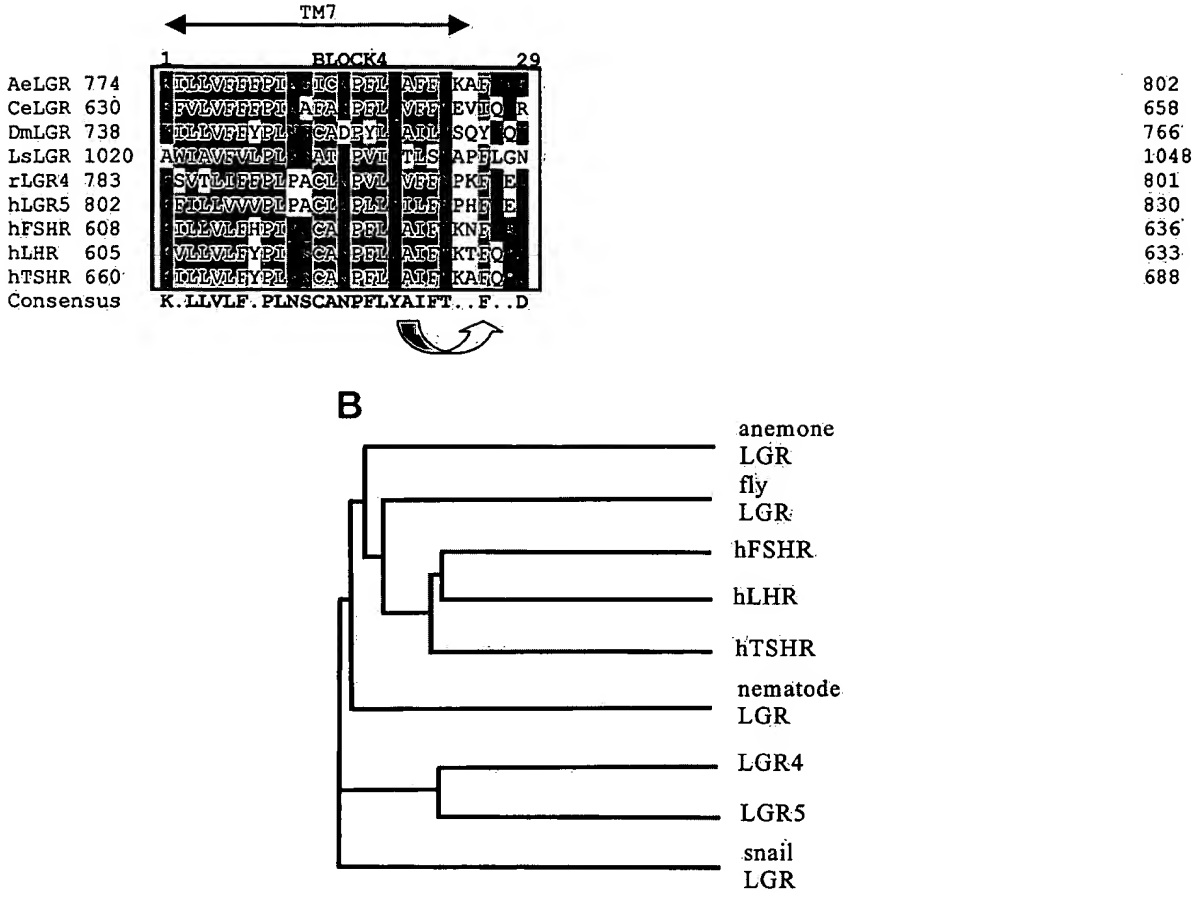


Fig. 4. Comparison of Protein Architectures of Mammalian LGRs and Their Relationship to Similar Receptors Found in Sea Anemone, Nematode, Fly, and Snail

A, Sequence comparison of the highly conserved TM regions of LGRs from diverse species [sea anemone: Ae, *Anthopleura elegantissima*; nematode: Ce, *Caenorhabditis elegans*; fly: Dm, *Drosophila melanogaster*; snail: Ls, *Lymnaea stagnalis*; rat (r) and human (h)]. Based on BLOCK Maker analysis, four highly conserved ungapped blocks were identified in the TM and flanking sequences of nine LGRs from diverse species. *Boxed areas* are sequence blocks with multiply aligned ungapped segments corresponding to



the most highly conserved regions of proteins. Secondary structure predictions (α-helices: *curved arrows*; β- strand: *straight arrow*) *below* the sequences were derived based on the PHD algorithm and DSC (60, 61). Chemically similar residues are *lightly shaded* whereas conserved charged residues are *heavily shaded*. Consensus residues represent identity among at least five of the nine receptors, B, Phylogenetic relatedness of LGRs from diverse species. Based on sequence comparison of the entire receptor proteins, LGRs can be divided into three subgroups: one containing the snail LGR, one containing mammalian LGR4 and LGR5, and a third one containing human gonadotropin and TSH receptors together with LGRs from fly, nematode, and sea anemone.

motifs and leucine-rich repeats together with the seven-TM region (31).

It is clear that the LGR genes encode a conserved subgroup of seven TM receptors of ancient origin. Based on the evolutionary relationship of LGRs in diverse species and the homologous amino acid sequences found in nine vertebrate and invertebrate LGRs, future searches could identify additional members of this subfamily of GPCRs in the mammalian genome. Because LGR4 and LGR5 appear to diverge from known gonadotropin and TSH receptors early during evolution before the formation of the pituitary gland and exhibit high similarity to LGRs found in lower vertebrates, they could subserve physiological functions associated with the primitive LGRs found in Cnidarians, one of the most primitive animals with a sensory system in the animal kingdom. A wide tissue distribution pattern of the mRNAs for these proteins further suggested that their physiological roles might be different from mammalian gonadotropin and TSH receptors known to play tissue-specific functions. The availability of cDNAs for LGR4 and LGR5 allows for future identification of their specific ligands by employing an anchored receptor approach found to be useful for the solubilization of the ligand-binding domains of glycoprotein hormone receptors (56) and for the elucidation of their physiological functions.

MATERIALS AND METHODS

Computational Analysis

Human sequences related to the sea anemone and *Drosophila* glycoprotein hormone receptors (12, 13) were identified from the EST database (dbEST) at the National Center for Biotechnology Information (NIH, Bethesda, MD) by using the BLAST server with the BLOSUM62 protein comparison matrix (57). The alignment of LGR ecto- and TM domain sequences was carried out by CLUSTALW (58); this program also calculated the branching order of aligned se-

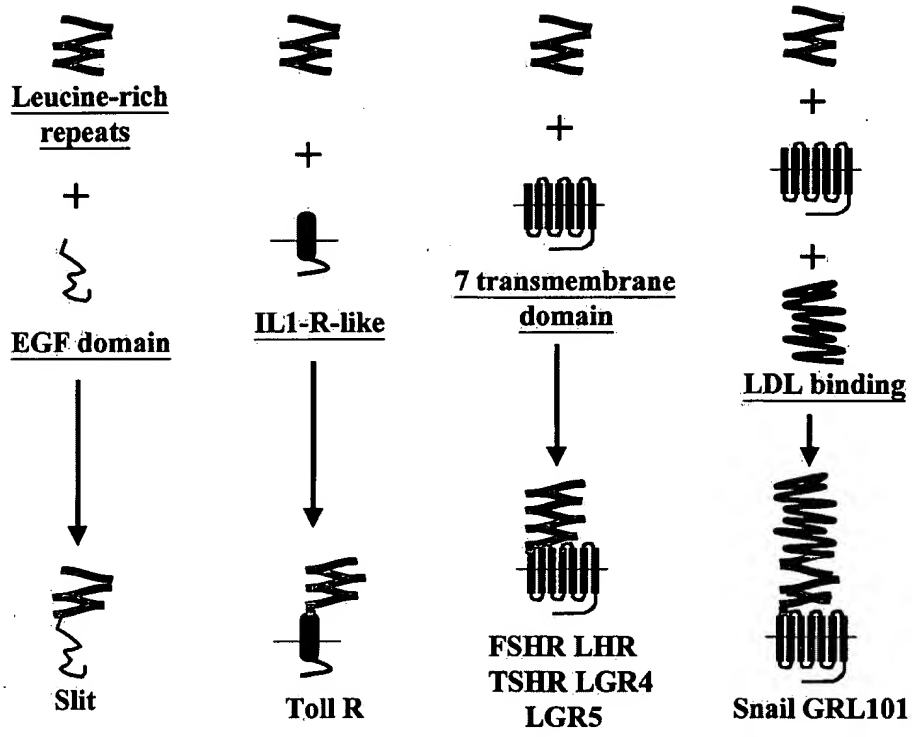


Fig. 5. Hypothetical Model on the Evolution of Diverse Genes Containing Leucine-Rich Repeats

Gene duplication and recombination probably account for the evolution of LGRs and related leucine-rich repeat-containing proteins. The motifs containing leucine-rich sequences or TM domains are the basic units to build modules of leucine-rich repeats and seven-TM helices through gene duplication. Fusion of ancestral leucine-rich repeat modules with epidermal growth factor (EGF) domain or interleukin (IL) 1 receptor (R)-like domain led to the formation of secreted slit protein and Toll receptors (R), respectively. In contrast, fusion of leucine-rich repeats with seven-TM modules led to the evolution of LGRs, whereas the fusion of leucine-rich repeats with seven-TM domains plus a low density lipoprotein (LDL)-binding region led to the formation of the GRL101 receptor found in snail.

quences by the neighbor-joining algorithm (10,000 bootstrap replications provided confidence values for the tree groupings). Conserved alignment patterns were drawn by the CONSENSUS program (Internet URL http://www.bork.embl-heidelberg.de/Alignment/consensus.html). The PRINTS library of protein fingerprints (http://www.biochem.ucl.ac.uk/bsm/dbbrowser/PRINTS/PRINTS.html) (18, 59) identified the myriad leucine-rich repeats present in the ectodomains of LGRs with a compound "Leurichrpt" motif that flexibly matches N- and C-terminal features of divergent leucine-rich repeats. The BLOCK Maker website (http://blocks.fhcrc.org) was used to align and generate the highly conserved ungapped blocks of the TM regions of diverse LGRs from vertebrates and invertebrates using both full-length receptor and TM segment sequences. The use of two different methods, Motif and Gibbs samplings, confirms for close relatedness. The blocks predicted from this alignment were then used to construct the neighbor-joining tree for the examination of possible subfamily relationships. Two algorithms whose three-state accuracy is greater than 72%, the neural network program PHD (60) and the statistical prediction method DSC (61; http://bonsai.lif.icnet.uk/dsc/manual.html), were used to derive a consensus secondary structure for the TM domain of different LGRs.

Cloning of Full-Length LGR4 and LGR5 cDNAs

Human ESTs showing high homology to two nonoverlapping regions of the gonadotropin receptors were identified. Clones AA312798 and AA298810 were found to encode TM 4 to TM 5 of the putative receptor (LGR4), whereas AA460529 and AA424098 encode TM 2 to TM 3 of another putative receptor (LGR5). Using these ESTs to further search the GenBank EST division database, overlapping EST sequences were aligned to obtain the longest ORF for each initial clone. Relevant EST clones were obtained from the I.M.A.G.E. consortium (info@image.llnl.gov) via Genome System, Inc. (St. Louis, MO).

Based on the longest human ORF, specific primers were designed for PCR amplification of LGR4 and LGR5 cDNA fragments from rat ovary and human placenta, respectively. After hybridization with labeled EST clones and confirmation of DNA sequences by dideoxy DNA sequencing, specific receptor fragments isolated were used to design primers to prepare sub-cDNA libraries enriched with specific receptor cDNAs. For 5′ extension, reverse transcription was performed using rat ovarian and human placenta mRNA preparations and receptor-specific primers. After second-strand synthesis, the enriched cDNA pool was tailed at 5′-ends with specific adaptor sequences to allow further PCR amplification. For 3′-extension, rat ovarian or human placenta mRNAs were reversed transcribed using oligo-dT, followed by second-strand synthesis using receptor-specific primers and adaptor tailing. These minilibraries were further used as templates for PCR amplification of upstream or downstream cDNAs specific for each receptor using internal primers. PCR products with a strong hybridization signal to each receptor cDNA fragment were subcloned into the pUC18 or pcDNA3 vectors. After screening of these sublibraries based on colony hybridization using specific receptor probes, clones with

5′- or 3′-sequences of the putative receptors were identified and isolated for DNA sequencing. As needed, the procedure was repeated up to three times to generate cDNAs encoding the complete ORF of each putative receptor for sequence analysis and for the expression of receptor proteins in eukaryotic cells. The entire coding sequences of each gene were also amplified with specific primers flanking the entire ORF in independent experiments. At least three independent PCR clones were sequenced to verify the authenticity of coding sequences.

Tissue Expression of LGR4 and LGR5 mRNAs

Human multiple tissue blots, containing ~2 μg of poly(A)+ RNA per lane, were purchased from CLONTECH (Palo Alto, CA; catalog number 7759, 7760, and 7767). Northern blot analyses were performed using tissue blots after hybridization of labeled receptor cDNA probes. Membranes were prehybridized for 1 h at 60 C in the ExpressHyb solution (CLONTECH). This was followed by hybridization under the same condition for 2 h but with 1×10^6 cpm/ml of ^{32}P-labeled LGR4, LGR5, or β-actin cDNA probe. After hybridization, the membranes were washed twice in 2 × saline-sodium citrate (SSC), 0.5% SDS at room temperature, followed by two washes in 0.2 × SSC, 0.5% SDS at 60 C before exposure to Kodak RX films (Eastman Kodak, Rochester, NY).

Construction of Chimeric Receptor cDNAs and Analysis of Signal Transduction and Ligand Binding

PCR-based mutagenesis was performed using overlapping primers to construct cDNAs for chimeric LH/LGR4 and LH/LGR5 receptors as described previously (29). L(EC)LGR4(TM) and L(EC)LGR5(TM) represent chimeric receptors with the ectodomain of human LH receptor and the TM and C-terminal tail of LGR4 or LGR5 with the junctional sequences of PEPDA-FKPCEYLLGS and PEPDA-FKPCEHLLDG, respectively. All cDNAs were subcloned into the expression vector pcDNA3 (Invitrogen, San Diego, CA). Both the fidelity of PCR-amplified regions and the junctional sequences were confirmed by DNA sequencing on both strands. 293 cells derived from human embryonic kidney fibroblast were maintained in DMEM/Ham's F-12 (Life Technologies, Inc., Gaithersburg, MD) supplemented with 10% FBS, 100 μg/ml penicillin, 100 μg/ml streptomycin, and 2 mM L-glutamine. The cells were transfected with receptor cDNAs as described (29) by the calcium phosphate precipitation method. Cells transfected with the empty plasmid (mock) served as negative controls. Cells were placed on 24-well tissue culture plates (Corning, Corning, NY) and preincubated at 37 C for 30 min in the presence of 0.25 mM 3-isobutyl-1-methyl xanthine (Sigma Chemical Co., St. Louis, MO) before treatment with or without hCG for 5 h. At the end of incubation, cells and medium in each well were frozen and thawed once and then heated to 95 C for 3 min to inactivate phosphodiesterase activity. Total cAMP was measured in triplicates by specific RIA. All experiments were repeated at least three times using cells from independent transfection. Statistical analysis was performed using Student's *t* test.

For ligand binding analysis of the chimeric receptors, hCG was iodinated by the lactoperoxidase method and characterized by radioligand receptor assay using human LH receptors stably expressed in 293 cells (29). Specific activity and maximal binding of the labeled hCG were 100,000–150,000 cpm/ng and 40–50%, respectively. To estimate ligand binding on the cell surface, cells were washed twice with PBS and collected in PBS before centrifugation at 400 × *g* for 5 min. Pellets were resuspended in PBS containing 0.1% BSA, and 200,000 cells/300 μl were incubated with a nearly saturating amount of labeled hCG at room temperature for 18–22 h in the presence or the absence of hCG. At the end of incubation, cells were centrifuged and washed twice with PBS. Radioactivities in the pellets were determined in a β-counter.

Genomic Analysis and Chromosomal Localization of LGR4 and LGR5

For studies on the conservation of LGR4 and LGR5 genes, the Zoo blots (CLONTECH) containing genomic DNA from different vertebrates were hybridized with ^{32}P-labeled rat LGR4 or human LGR5 cDNA probe under moderate stringency conditions.

To isolate genomic clones for LGR4 and LGR5, several genomic DNA fragments were isolated from a human bacterial artificial chromosome (BAC) genomic DNA library (Genome Systems, Inc.) using the near full-length LGR4 or LGR5 cDNA probes. The genomic fragments were then confirmed by Southern blot hybridization. For the identification of the chromosomal localization of LGR4 and LGR5 genes, genomic fragments (>100 Kb) of LGR4 and LGR5 were used as probes for FISH to human metaphase chromosomes (SeeDNA Biotech, Inc., Toronto, Ontario, Canada). Denatured chromosomes from synchronous cultures of human lymphocytes were hybridized with biotinylated probes for signal localization.

Acknowledgments

We thank C. Spencer for editorial assistance.

Received July 21, 1998. Re-revision received September 9, 1998. Accepted September 11, 1998.

Address requests for reprints to: Dr. A. J. W. Hsueh, Division of Reproductive Biology, Department of Gynecology/Obstetrics, Stanford University School of Medicine, 300 Pasteur Drive, Stanford, California 94305-5317. E-mail: aaron.hsueh@stanford.edu.

This work was supported by NIH Grant HD-23273.

*The first two authors contributed equally to this work. S.L. is on leave from the Department of Obstetrics and Gynecology, Teikyo University, Tokyo, Japan.

REFERENCES

1. Roth NS, Lefkowitz RJ, Caron MG 1991 Structure and function of the adrenergic receptor family. Adv Exp Med Biol 308:223–238
2. Strader CD, Fong TM, Tota MR, Underwood D, Dixon RA 1994 Structure and function of G protein-coupled receptors. Annu Rev Biochem 63:101–132
3. McFarland KC, Sprengel R, Phillips HS, Kohler M, Rosemblit N, Nikolics K, Segaloff DL, Seeburg PH 1989 Lutropin-choriogonadotropin receptor: an unusual member of the G protein-coupled receptor family. Science 245:494–499
4. Vassart G, Dumont JE 1992 The thyrotropin receptor and the regulation of thyrocyte function and growth. Endocr Rev 13:596–611
5. Sprengel R, Braun T, Nikolics K, Segaloff DL, Seeburg PH 1990 The testicular receptor for follicle stimulating hormone: structure and functional expression of cloned cDNA. Mol Endocrinol 4:525–530
6. Gutkind JS 1998 The pathways connecting G protein-coupled receptors to the nucleus through divergent mitogen-activated protein kinase cascades. J Biol Chem 273:1839–1842
7. Selbie LA, Hill SJ 1998 G protein-coupled-receptor cross-talk: the fine-tuning of multiple receptor-signalling pathways. Trends Pharmacol Sci 19:87–93

8. Ji TH, Ryu KS, Gilchrist R, Ji I 1997 Interaction, signal generation, signal divergence, and signal transduction of LH/CG and the receptor. Recent Prog Horm Res 52:431–453
9. Dufau ML 1998 The lutenizing hormone receptor. Annu Rev Physiol 60:461–496
10. Kohn LD, Shimura H, Shimura Y, Hidaka A, Giuliani C, Napolitano G, Ohmori M, Laglia G, Saji M 1995 The thyrotropin receptor. Vitam Horm 50:287–384
11. Simoni M, Gromoll J, Nieschlag E 1997 The follicle-stimulating hormone receptor: biochemistry, molecular biology, physiology, and pathophysiology. Endocr Rev 18:739–773
12. Hauser F, Nothacker H, Grimmelikhuijzen CJP 1997 Molecular cloning, genomic organization, and developmental regulation of a novel receptor from *Drosophila melanogaster*, structurally related to members of the thyroid-stimulating hormone, follicle-stimulating hormone, luteinizing hormone/choriogonadotropin receptor family from mammals. J Biol Chem 272:1002–1010
13. Nothacker HP, Grimmelikhuijzen CJ 1993 Molecular cloning of a novel, putative G protein-coupled receptor from sea anemones structurally related to members of the FSH, TSH, LH/CG receptor family from mammals. Biochem Biophys Res Commun 197:1062–1069
14. Lee CY, Rechler MM 1995 Purified rat acid-labile subunit and recombinant human insulin-like growth factor (IGF)-binding protein-3 can form a 150-kilodalton binary complex *in vitro* in the absence of IGFs. Endocrinology 136:4982–4989
15. Leong SR, Baxter RC, Camerato T, Dai J, Wood WI 1992 Structure and functional expression of the acid-labile subunit of the insulin-like growth factor-binding protein complex. Mol Endocrinol 6:870–876
16. Twigg SM, Baxter RC 1998 Insulin-like growth factor (IGF)-binding protein 5 forms an alternative ternary complex with IGFs and the acid-labile subunit. J Biol Chem 273:6074–6079
17. Rothberg JM, Jacobs JR, Goodman CS, Artavanis-Tsakonas S 1990 Slit: an extracellular protein necessary for development of midline glia and commissural axon pathways contains both EGF and LRR domains. Genes Dev 4:2169–2187
18. Buchanan SG, Gay NJ 1996 Structural and functional diversity in the leucine-rich repeat family of proteins. Prog Biophys Mol Biol 65:1–44
19. Rock FL, Hardiman G, Timans JC, Kastelein RA, Bazan JF 1998 A family of human receptors structurally related to *Drosophila* Toll. Proc Natl Acad Sci USA 95:588–593
20. Kajava AV, Vassart G, Wodak SJ 1995 Modeling of the three-dimensional structure of proteins with the typical leucine-rich repeats. Structure 3:867–877
21. Kajava AV 1998 Structural diversity of leucine-rich repeat proteins. J Mol Biol 277:519–527
22. Parmentier M, Libert F, Maenhaut C, Lefort A, Gerard C, Perret J, Van Sande J, Dumont JE, Vassart G 1989 Molecular cloning of the thyrotropin receptor. Science 246:1620–1622
23. Luttrell LM, Ostrowski J, Cotecchia S, Kendall H, Lefkowitz RJ 1993 Antagonism of catecholamine receptor signaling by expression of cytoplasmic domains of the receptors. Science 259:1453–1457
24. Huang J, Puett D 1995 Identification of two amino acid residues on the extracellular domain of the lutropin/choriogonadotropin receptor important in signaling. J Biol Chem 270:30023–30028
25. Gudermann T, Nurnberg B, Schultz G 1995 Receptors and G proteins as primary components of transmembrane signal transduction. I. G-protein-coupled receptors: structure and function. J Mol Med 73:51–63
26. Hong S, Ryu KS, Oh MS, Ji I, Ji TH 1997 Roles of transmembrane prolines and proline-induced kinks of the lutropin/choriogonadotropin receptor. J Biol Chem 272:4166–4171
27. Daaka Y, Luttrell LM, Lefkowitz RJ 1997 Switching of the coupling of the beta2-adrenergic receptor to different G proteins by protein kinase A. Nature 390:88–91
28. Pawson T, Scott JD 1997 Signaling through scaffold, anchoring, and adaptor proteins. Science 278: 2075–2080
29. Kudo M, Osuga Y, Kobilka BK, Hsueh AJW 1996 Transmembrane regions V and VI of the human luteinizing hormone receptor are required for constitutive activation by a mutation in the third intracellular loop. J Biol Chem 271:22470–22478
30. Wilson R, Ainscough R, Anderson K, Baynes C, Berks M, Bonfield J, Burton J, Connell M, Copsey T, *et al.* 1994 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*. Nature 368:32–38
31. Tensen CP, Van Kesteren ER, Planta RJ, Cox KJ, Burke JF, van Heerikhuizen H, Vreugdenhil EA 1994 G protein-coupled receptor with low density lipoprotein-binding motifs suggests a role for lipoproteins in G protein-linked signal transduction. Proc Natl Acad Sci USA 91:4816–4820
32. Henikoff S, Henikoff JG, Alford WJ, Pietrokovski S 1995 Automated construction and graphical presentation of protein blocks from unaligned sequences. Gene 163:GC17–GC26
33. Ryu K, Lee H, Kim S, Beauchamp J, Tung CS, Isaacs NW, Ji I, Ji TH 1998 Modulation of high affinity hormone binding. Human choriogonadotropin binding to the exodomain of the receptor is influenced by exoloop 2 of the receptor. J Biol Chem 273:6285–6291
34. Gromoll J, Ried T, Holtgreve-Grez H, Nieschlag E, Gudermann T 1994 Localization of the human FSH receptor to chromosome 2 p21 using a genomic probe comprising exon 10. J Mol Endocrinol 12:265–271
35. Kobe B, Deisenhofer J 1993 Crystal structure of porcine ribonuclease inhibitor, a protein with leucine-rich repeats. Nature 366:751–756
36. Medzhitov R, Preston-Hurlburt P, Janeway Jr CA 1997 A human homologue of the *Drosophila* Toll protein signals activation of adaptive immunity. Nature 388:394–397
37. Williams MJ, Rodriguez A, Kimbrell DA, Eldon ED 1997 The 18-wheeler mutation reveals complex antibacterial gene regulation in *Drosophila* host defense. EMBO J 16:6120–6130
38. Jiang X, Dreano M, Buckler DR, Cheng S, Ythier A, Wu H, Hendrickson WA, el Tayar N 1995 Structural predictions for the ligand-binding region of glycoprotein hormone receptors and the nature of hormone-receptor interactions. Structure 3:1341–1353
39. Moyle WR, Campbell RK, Rao SN, Ayad NG, Bernard MP, Han Y, Wang, Y 1995 Model of human chorionic gonadotropin and lutropin receptor interaction that explains signal transduction of the glycoprotein hormones. J Biol Chem 270:20020–20031
40. Rajaram S, Baylink DJ, Mohan S 1997 Insulin-like growth factor-binding proteins in serum and other biological fluids: regulation and functions. Endocr Rev 18:801–831
41. Takeuchi Y, Kodama Y, Matsumoto T 1994 Bone matrix decorin binds transforming growth factor-beta and enhances its bioactivity. J Biol Chem 269:32634–32638
42. Morisato D, Anderson KV 1994 Signaling pathways that establish the dorsal-ventral pattern of the *Drosophila* embryo. Cell 76:677–688
43. Ehnis T, Dieterich W, Bauer M, Kresse H, Schuppan D 1997 Localization of a binding site for the proteoglycan decorin on collagen XIV (undulin). J Biol Chem 272:20414–20419
44. Casanova J, Furriols M, McCormick CA, Struhl G 1995 Similarities between trunk and spatzle, putative extracellular ligands specifying body pattern in *Drosophila*. Genes Dev 9:2539–2544
45. Aittomaki K, Lucena JL, Pakarinen P, Sistonen P, Tap-

anainen J, Gromoll J, Kaskikari R, Sankila EM, Lehvaslaiho H, Engel AR, et al. 1995 Mutation in the follicle-stimulating hormone receptor gene causes hereditary hypergonadotropic ovarian failure. Cell 82:959–968
46. Tapanainen JS, Aittomaki K, Min J, Vaskivuo T, Huhtaniemi IT 1997 Men homozygous for an inactivating mutation of the follicle-stimulating hormone (FSH) receptor gene present variable suppression of spermatogenesis and fertility. Nat Genet 15:205–206
47. Puett D, Bhowmick N, Fernandez LM, Huang J, Wu C, Narayan P 1996 hCG-receptor binding and transmembrane signaling. Mol Cell Endocrinol 125:55–64
48. Osuga Y, Hayashi M, Kudo M, Conti M, Kobilka B, Hsueh AJW 1997 Co-expression of defective luteinizing hormone receptor fragments partially reconstitutes ligand-induced signal generation. J Biol Chem 272: 25006–25012
49. Baldwin JM 1993 The probable arrangement of the helices in G protein-coupled receptors. EMBO J 12:1693–1703
50. Hoflack J, Hibert MF, Trumpp-Kallmeyer S, Bidart JM 1993 Three-dimensional models of gonado-thyrotropin hormone receptor transmembrane domain. Drug Des Discov 10:157–171
51. Lin Z, Shenker A, Pearlstein R 1997 A model of the lutropin/choriogonadotropin receptor: insights into the structural and functional effects of constitutively activating mutations. Protein Eng 10:501–510
52. McDonald T, Wang R, Bailey W, Xie G, Chen F, Caskey CT, Liu Q 1998 Identification and cloning of an orphan G protein-coupled receptor of the glycoprotein hormone receptor subfamily. Biochem Biophys Res Commun 247:266–270
53. Zhu SZ, Wang SZ, Hu J, el-Fakahany EE 1994 An arginine residue conserved in most G protein-coupled receptors is essential for the function of the m1 muscarinic receptor. Mol Pharmacol 45:517–523
54. Henikoff S, Greene EA, Pietrokovski S, Bork P, Attwood TK, Hood L 1997 Gene families: the taxonomy of protein paralogs and chimeras. Science 278:609–614
55. Tsai-Morris CH, Buczko E, Wang W, Xie XZ, Dufau ML 1991 Structural organization of the rat luteinizing hormone (LH) receptor gene. J Biol Chem 266:11355–11359
56. Osuga Y, Kudo M, Kaipia A, Kobilka B, Hsueh AJW 1997 Derivation of functional antagonists using N-terminal extracellular domain of gonadotropin and thyrotropin receptors. Mol Endocrinol 11:1659–1668
57. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ 1997 Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402
58. Thompson JD, Higgins DG, Gibson TJ, Thompson JD, Higgins DG, Gibson TJ 1994 CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680
59. Attwood TK, Beck ME, Bleasby AJ, Degtyarenko K, Michie AD, Parry-Smith DJ 1997 Novel developments with the PRINTS protein fingerprint database. Nucleic Acids Res 25:212–217
60. Rost B, Sander C 1996 Bridging the protein sequence-structure gap by structure predictions. Annu Rev Biophys Biomol Struct 25:113–136
61. King RD, Sternberg MJ 1996 Identification and application of the concepts important for accurate and reliable protein secondary structure prediction. Protein Sci 5:2298–2310
62. Chang Z, Price BD, Bockheim S, Boedigheimer MJ, Smith R, Laughon A 1993 Molecular and genetic characterization of the *Drosophila* tartan gene. Dev Biol 160:35–332
63. Minegishi T, Nakamura K, Takakura Y, Ibuki Y, Igarashi M 1991 Cloning and sequencing of human FSH receptor cDNA. Biochem Biophys Res Commun 175:1125–1130
64. Jia XC, Oikawa M, Bo M, Tanaka T, Ny T, Boime I, Hsueh AJW 1991 Expression of human luteinizing hormone (LH) receptor: interaction with LH and chorionic gonadotropin from human but not equine, rat, and ovine species. Mol Endocrinol 5:759–768